

**WEST**[Help](#)[Logout](#)[Interrupt](#)[Main Menu](#) [Search Form](#) [Posting Counts](#) [Show S Numbers](#) [Edit S Numbers](#) [Preferences](#)**Search Results -**

Term	Documents
TTR.USPT.	474
TTRS.USPT.	14
PROMOTER.USPT.	40068
PROMOTERS.USPT.	31139
((TTR ADJ PROMOTER) AND 1).USPT.	2

Database: 

US Patents Full-Text Database

US Pre-Grant Publication Full-Text Database

JPO Abstracts Database

EPO Abstracts Database

Derwent World Patents Index

IBM Technical Disclosure Bulletins

Refine Search:

11 and ttr promoter

[Clear](#)**Search History**

Today's Date: 1/9/2002



<u>DB Name</u>	<u>Query</u>	<u>Hit Count</u>	<u>Set Name</u>
USPT	11 and ttr promoter	2	<u>L13</u>
USPT	11 and transthyretin gene promoter	0	<u>L12</u>
USPT	11 and hnf 3 albumin promoter	2	<u>L11</u>
USPT	11 and hnf3 albumin promoter	0	<u>L10</u>
USPT,PGPB,JPAB,EPAB,DWPI	13 and hnf-3 albumin promoter	3	<u>L9</u>
USPT,PGPB,JPAB,EPAB,DWPI	13 and ttr gene promoter	3	<u>L8</u>
USPT,PGPB,JPAB,EPAB,DWPI	13 and hnf3 albumin promoter	0	<u>L7</u>
USPT,PGPB,JPAB,EPAB,DWPI	15 and ser adj phe	8	<u>L6</u>
USPT,PGPB,JPAB,EPAB,DWPI	14 and junction	15	<u>L5</u>
USPT,PGPB,JPAB,EPAB,DWPI	13 and heavy chain and light chain	23	<u>L4</u>
USPT,PGPB,JPAB,EPAB,DWPI	12 and factor VIII	82	<u>L3</u>
USPT,PGPB,JPAB,EPAB,DWPI	11 and pharmaceutical composition	903	<u>L2</u>
USPT,PGPB,JPAB,EPAB,DWPI	adeno associated virus or aav	2336	<u>L1</u>



11-11-11



Wed Jan 9 14:38:15 2002

US-09-740-211-13-copy\_11000\_11933.rge

Page 1

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: January 8, 2002, 17:16:51 : Search time 7258.17 seconds  
(without alignments)  
2122.900 Million cell updates/sec

Title: US-09-740-211-13-copy\_11000\_11933  
Sequence: 1 aaacggcagagagtgtag.....cttaagctgcgcgcgcag 934

Scoring table: IDENTITY\_NUC  
Gapop 10.0 / Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues 2944260

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 08  
Maximum Match 1008  
Listing first 45 summaries

Database:

GenEmbl:  
1: gb\_ba:  
2: gb\_hg:  
3: gb\_in:  
4: gb\_om:  
5: gb\_ov:  
6: gb\_pat:  
7: gb\_ph:  
8: gb\_pl:  
9: gb\_pr:  
10: gb\_ro:  
11: gb\_sts:  
12: gb\_sy:  
13: gb\_un:  
14: gb\_vi:  
15: em\_ba:  
16: em\_fun:  
17: em\_hum:  
18: em\_in:  
19: em\_om:  
20: em\_or:  
21: em\_ov:  
22: em\_pat:  
23: em\_ph:  
24: em\_pi:  
25: em\_ro:  
26: em\_sts:  
27: em\_sy:  
28: em\_un:  
29: em\_vi:  
30: em\_hgo\_hum:  
31: em\_hgo\_inv:  
32: em\_hgo\_rod:  
33: em\_hg\_hum:  
34: em\_hg\_inv:  
35: em\_hg\_rod:  
36: em\_hg\_other:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	934	100.0	11933	ARI38377	ARI38377 Sequence
2	934	100.0	11933	ARI46887	ARI46887 Sequence
3	923.2	98.8	42529	CVU39284	U39284 Cloning vec
4	923.2	98.8	42530	CVU39285	U39285 Cloning vec
5	923.2	98.8	42531	CVU39286	U39286 Cloning vec
6	923.2	98.8	42531	CVU39286	U39286 Cloning vec
7	923.2	98.8	42531	CVU39286	U39286 Cloning vec
8	921.6	98.7	9205	XXU02427	U02427 Cloning vec
9	921.6	98.7	9205	XXU02427	U02427 Cloning vec
10	657.4	70.4	751	LAMOPR	M38294 Bacteriophage
11	448.2	48.0	805	STHK0220	X60309 Coliphage H
12	436.2	46.7	40751	AF069308	AF069308 Bacteriophage
13	436.2	46.7	40751	AF069308	AF069308 Bacteriophage
14	429.8	45.7	38297	AF069358	AF069358 Bacteriophage
15	426.8	45.5	6709	BP0011580	AF0011580 Bacteriophage
16	425.8	43.9	10390	AE005492	AE005492 Escherich
17	409.8	43.9	270365	POP22NIN	AF217253 Bacteriophage
18	401	42.9	41724	AC009667	AC009667 Homo sapi
19	401	42.9	41724	AC009667	AC009667 Homo sapi
20	401	42.9	41724	AC009667	AC009667 Homo sapi
21	125.8	13.5	191683	AE005288	AE005288 Escherich
22	58	6.2	222605	AE005555	AE005555 Escherich
23	58	6.2	222605	AE005555	AE005555 Escherich
24	48.4	5.2	17300	D90800	D90800 E. coli geno
25	48.4	5.2	17300	D90800	D90800 E. coli geno
26	48.4	5.2	17300	D90800	D90800 E. coli geno
27	48.4	5.2	17300	D90800	D90800 E. coli geno
28	48.4	5.2	17300	D90800	D90800 E. coli geno
29	42.2	4.5	137956	AC010918	AC010918 Homo sapi
30	42.2	4.5	137956	AC010918	AC010918 Homo sapi
31	42.2	4.5	137956	AC010918	AC010918 Homo sapi
32	42.2	4.5	137956	AC010918	AC010918 Homo sapi
33	42.2	4.5	137956	AC010918	AC010918 Homo sapi
34	42.2	4.5	137956	AC010918	AC010918 Homo sapi
35	42.2	4.5	137956	AC010918	AC010918 Homo sapi
36	42.2	4.5	137956	AC010918	AC010918 Homo sapi
37	35.6	3.8	156555	AL157390	AL157390 Homo sapi
38	35.6	3.8	156555	AL157390	AL157390 Homo sapi
39	35.6	3.8	156555	AL157390	AL157390 Homo sapi
40	35.6	3.8	156555	AL157390	AL157390 Homo sapi
41	35.6	3.8	156555	AL157390	AL157390 Homo sapi
42	35.4	3.8	163095	AP002523	AP002523 Oryza sat
43	35.4	3.8	163095	AP002523	AP002523 Oryza sat
44	35.4	3.8	163095	AP002523	AP002523 Oryza sat
45	35.2	3.8	324484	AE003584	AE003584 Drosophila

ALIGNMENTS

RESULT 1  
LOCUS ARI38377 11933 bp DNA  
DEFINITION Sequence 13 from patent US 6200560.  
ACCESSION ARI38377  
VERSION ARI38377.1 GI:14480722  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 11933)  
AUTHORS Couto, L.B., Colosi, P.C. and Qian, X.  
TITLE Adeno-associated virus vectors for expression of factor VIII by target cells  
JOURNAL Patent: US 6200560-A 13 MAR-2001;  
FEATURES  
1. 11933 Location/Qualifiers  
BASE COUNT 3258 a 2818 c 2717 g 3140 t

PAT 16-JUN-2001

Query Match 100.0% Score 934; DB 6; Length 11933;  
Best Local Similarity 100.0%; Pred. No. 1.5e-286;  
Matches 934; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 aaagcgaggaaggtttgttagcgacacccctcccccagcgttttaagaagtcgtgta 60
Db 11000 AAAGCGAGGAGGTTGTAGCGCGACCTCTGCCACCCGCTTTCACCAAGTCATGTGTA 11059
OY 61 aaagcgaggaaggtttgttagcgacacccctcccccagcgttttaagaagtcgtgta 120
Db 11060 AAAGCGCGAGCGTAACTATTAATGAATGACGACAGACAGTGGCTACGGCTCAGTT 11119
OY 121 tgggttgctgttctgttgagcgagcgtgacgctgtagcagatttggttcggcttcgc 180
Db 11120 TGGGTTGGCTGTGCTGCTGAGCGGATGACGCTGTACCATTTGGTGTGCTGCTGCT 11179
OY 181 ttcggatttcgcttaattcagcaacaaggaagagacattgctaacagagctcgccgac 240
Db 11180 TTCCGGATTTCGCTTAATTACGACCAACGAAAGACACTGGCTAACCGAGCTCGCCGAC 11239
OY 241 tcttcagcattatcgactcaattccttcttctgttctgtgcagataaaatcccgaaac 300
Db 11240 TCTTCACGATTATCGACTCAATCTCTTACCTGTGTGTCAGATATTAATAATCCCAAAAC 11299
OY 301 cgtatgcagcgtctactacttaactcgtgacgtgttcggagattgcatltagacacc 360
Db 11300 CGTTATGACAGGCTCTACTACTATTAACCTGCGAACTGTTTCGGGATTGCAATTTGACAGCTTC 11359
OY 361 tctgctgcatgtgttgagttcagcaagacgtgagtgacaaactagacggagtcg 420
Db 11360 TCTGCTGCGATGTGTGAGTTCAGACGATACGTGAAAGTGCACAACTAGCGGAACTCG 11419
OY 421 gtaagtaagcgcgcctctcttctactctacacacacagagcaatlaaccatcgttga 480
Db 11420 GTAGTAAGCGCGCCTCTCTTCTCTCTCTACACACAGCAAGCAATTAACCATCTGTGA 11479
OY 481 gtcaaatltaaccaatttatcacaataagtcacatcgtccgttaataatgttgcacac 540
Db 11480 GTCAAAATTTACCAAAATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 11539
OY 541 gtggcaatcactgctctgaacgttgacagcaatlaaaatgttctgagatgaactc 600
Db 11540 GTGGCAATCACTGCTCTGATGACATGTCACAGCGCATATGCTGCGCTTGGCATCTTC 11599
OY 601 ttgttgacatgacacacacagagcgtacacagcgttlaacagttcgtgacagagtc 660
Db 11600 TTTGTGCGATGTCACACACAGCGCATATGCTGCGCTTGGCATCTTC 11659
OY 661 tctggttaaggtttggattgaacatcgtcagcagcgtatgctgctgctgacatc 720
Db 11660 TTGGTGAAGGTTTGGATTGAGCATGTCACAGCGCATATGCTGCGCTTGGCATCTTC 11719
OY 721 tgaatagcagcagccttctgactcttcgacactcttctgacaaactcccccagaatc 780
Db 11720 TGAATAGCGCGCCTTTCGATCTTTCGCAACTCTTCCGCCAGACCTC 11779
OY 781 tgtttgcaataatcaacgcagcagcgttgaacatcgtcagcagcgttgaacttc 840
Db 11780 TGTTTGCAATATCAACCGCAGCGCTGTACATGGAATCTGTGATCTTCCGCCGCG 11839
OY 841 cgtcgcgagcactaagcgaataatccgaataagcaatgttgcagagcacttgcaactc 900
Db 11840 CGTCGCGGACACATGCGCAATATGCGCATTAAGCGAATGTTGCGACGACTTGCAGTACTCT 11899
OY 901 tgccttaagtaattccttgaactcgtccctcgag 934
Db 11900 TGCCTTAAGTATTCCTTCAAGCTGCCCCCTCAGG 11933

```

RESULT 2  
AR146887

LOCUS AR146887 11933 bp DNA  
DEFINITION Sequence 13 from patent US 6221349.  
ACCESSION AR146887  
VERSION AR146887.1 GI:15110690  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 11933)  
AUTHORS Couto, L.B., Colosi, P.C. and Qian, X.  
TITLE Adeno-associated vectors for expression of factor VIII by target cells  
JOURNAL Patent: US 6221349-A 13 24-Apr-2001;  
FEATURES  
source location/Qualifiers  
BASE COUNT 3258 a 2818 c 2717 g 3140 t  
ORIGIN

Query Match 100.0% Score 934; DB 6; Length 11933;  
Best Local Similarity 100.0%; Pred. No. 1.5e-286;  
Matches 934; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 aaagcgaggaaggtttgttagcgacacccctcccccagcgttttaagaagtcgtgta 60
Db 11000 AAAGCGAGGAGGTTGTAGCGCGACCTCTGCCACCCGCTTTCACCAAGTCATGTGTA 11059
OY 61 aaagcgaggaaggtttgttagcgacacccctcccccagcgttttaagaagtcgtgta 120
Db 11060 AAAGCGCGAGCGTAACTATTAATGAATGACGACAGACAGTGGCTACGGCTCAGTT 11119
OY 121 tgggttgctgttctgttgagcgagcgtgacgctgtagcagatttggttcggcttcgc 180
Db 11120 TGGGTTGGCTGTGCTGAGCGGATGACGCTGTACCATTTGGTGTGCTGCTGCT 11179
OY 181 ttcggatttcgcttaattcagcaacaaggaagagacattgctaacagagctcgccgac 240
Db 11180 TTCCGGATTTCGCTTAATTACGACCAACGAAAGACACTGGCTAACCGAGCTCGCCGAC 11239
OY 241 tcttcagcattatcgactcaattccttcttctgttctgtgcagataaaatcccgaaac 300
Db 11240 TCTTCACGATTATCGACTCAATCTCTTACCTGTGTGTCAGATATTAATAATCCCAAAAC 11299
OY 301 cgtatgcagcgtctactacttaactcgtgacgtgttcggagattgcatltagacacc 360
Db 11300 CGTTATGACAGGCTCTACTACTATTAACCTGCGAACTGTTTCGGGATTGCAATTTGACAGCTTC 11359
OY 361 tctgctgcatgtgttgagttcagcaagacgtgagtgacaaactagacggagtcg 420
Db 11360 TCTGCTGCGATGTGTGAGTTCAGACGATACGTGAAAGTGCACAACTAGCGGAACTCG 11419
OY 421 gtaagtaagcgcgcctctcttctactctacacacacagagcaatlaaccatcgttga 480
Db 11420 GTAGTAAGCGCGCCTCTCTTCTCTCTCTACACACAGCAAGCAATTAACCATCTGTGA 11479
OY 481 gtcaaatltaaccaatttatcacaataagtcacatcgtccgttaataatgttgcacac 540
Db 11480 GTCAAAATTTACCAAAATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 11539
OY 541 gtggcaatcactgctctgaacgttgacagcaatlaaaatgttctgagatgaactc 600
Db 11540 GTGGCAATCACTGCTCTGATGACATGTCACAGCGCATATGCTGCGCTTGGCATCTTC 11599
OY 601 ttgttgacatgacacacacagagcgtacacagcgttlaacagttcgtgacagagtc 660
Db 11600 TTTGTGCGATTGCACACAGCGCTGTACATGGAATCTGTGATCTTCCGCCGCG 11659
OY 661 tctggttaaggtttggattgaacatcgtcagcagcgtatgctgctgacatc 720
Db 11660 TTGGTGAAGGTTTGGATTGAGCATGTCACAGCGCATATGCTGCGCTTGGCATCTTC 11719
OY 721 tgaatagcagcagccttctgactcttcgacactcttctgacaaactcccccagaatc 780

```

SLNGWRMAMFPABEAVPESLTCEDLPEADTVVVPWMQNHGTDAITNTNTVPTLVN  
PPEVPEPNTGCGSLITFNVNESHLOGGORTIFEDDVNAPMSGTEFRDVLHKKPTQIDSDH  
SEFDDLSAFIRAGENLAWMLKMSDSTYLEDDMGTOVASTAPFGGEIIDEKG  
VAREEDDSRAVLAEVOMCGELRDLRYTSLMGTOVASTAPFGGEIIDEKADYFREVRLNGL  
YABRPLTLNVENPGLTMSAEIPNLRYAVVELHTAGCTLIIEACDVYFREVRLNGL  
LNRKPLTLNRHREHNPILHGOVMOEQTVODILMKONNVPNRCSHPNHPMLT  
LCRGKGLYVYDANRITETGVAMPNRRLTDEGGRMLPASEEYTPRMVORNRHPSYIWSL  
LNGEKGACANDALRYMKISVDPSPRPVYEDSGMADVTADITICPNYALVDGDFPPAVP  
GNEGHSKANDALRYMKISVDPSPRPVYEDSGMADVTADITICPNYALVDGDFPPAVP  
KMSIKKMSLIPGTRPLITCEYAHAMGNSLGZGFAKYWOAFROYPLHOGGVYMOVVOG  
LITKDEGNSWMSVYGDREDFPNDPNOCRNMGVLFAFARTRHPPLTAKHDOOPFOPRS  
GOTIEVSESLYFPHSDNELHMMVALDSKPLASEVPLDVAPOGKQILELDELPOPS  
AGOLMIRVYVOCNPATAMEAGHISAMQOQVLEENSTYLPASASHIPLHLSMDPC  
IBLKACRMVORNSQFSLQMWIGDKKQILLPLPROCTRAEYPLNDIGSEARIDIPNM  
VERMKACHVOAEALLOCTADTLADAVLITVTHANMOHGKTLTSRKYTYIIDSQIM  
ATTVEVYASPTSPHARIGLOAQVAREYVWNLGLSPOMNPRDPLTACEDRDPLD  
SDMYTPVYFESNLTGCFRPELNYGHOMRQDFGNISRTSOOLMESHRLHLHAE  
GTWLNIDGFHMGIGDDDSWSPSYSAFOLSGAGRYHOLVWCOR  
complement(23874..23957)  
"note="multiple cloning site"  
complement(23973..24141)  
/note="multiple cloning site"  
24475..25665  
/gene="tetR"  
24475..25665  
/gene="tetR"

	Query Match	98.8%;	Score 923.2;	DB 12;	Length 42529;	
	Best Local Similarity	99.7%;	Pred. No. 4,4e-283;			
	Matches 925; Conservative	0;	Mismatches 3;	Indels 0;	Gaps 0	
QY	1 aaacggcagagagtgtgtagcggaacctctctgcaccgcgcttcacgaagttcatgtgta	60				
Db	35091 AAACGGCAGGAGGTGGTAGCGCACCTCTCCGCCACCCTCCGTTTCACAGAAGTCAATGTGA	39032				
QY	61 aaaggccgacgataactatctaagaatagaagcacagatgtgtaagtcgaagtt	120				
Db	39031 AAAGGCCGACGAGTACTATTACTAATAATTCAGACAGACAGTGCTTAACGGCTCATAGTT	38972				
QY	121 tggagtgtgctgttgccttgtagcgagatagcaagcctgtacaagatttgtgtgatccggtlctgc	180				
Db	38971 TGGGTTTGCTGTGCTCTGGCGGCAGTACGCGCTTCAGCATTTGTGTGATCCGGTTCCTGC	38912				
QY	181 ttccgcatlctgcgttaattcagcaaacagaaagacacitgycataccaggtctgcgcgac	240				
Db	38911 TTCCGGTATTCGCTTAATTCAGACACAACGAAAGACACTGGCTPAACCAAGCTGCCGAC	38852				
QY	241 tcctacagatcatcagcataatgctcttaacctgttltgagatatanaaatcccgaac	300				
Db	38851 TCCTACGATTATGCATCAATGCTTTACTCTGTGTGCAGATATAAAAAATCCCGAAC	38792				
QY	301 cgttatcgaagctctaactatactatctcgcagactgttctcgggttlitgattlltgcagactc	360				
Db	38791 CGTTATGCAAGGCTCTTAATACTTACTCGGAACCTGTTTCGGGATGTGCAATTTTCAGACCTC	38732				
QY	361 tctcgtctgcagatgttgaagttccaagaagatacgtcgaatlgaaccaactaagcgcgaatcg	420				

Db	38731	TCGGCTGGATGTTGGAGTTCAGACAGATACCTGCAACGACCAACTAGGGGGAAATG	38672
Qy	421	gagagaagcgcgcgcctctttcatctcaacacacagagaggaatlaacccatcggtga	38672
Db	38671	GTAGTAACCGCGGCTCTTTTCATCTCATTACCAACACAGCAATTAACCATGCTTGA	480
Qy	481	gtcaaatltaccacatlttltcaatlaaagaatcatatcgtcggtlaataatgttgcaccc	38612
Db	38611	GTCAAAATTTACCAATTTTATTCATTAATGATATCATGCGGTTAATATATGTTCATCC	540
Qy	541	gtggcaatcgtctgtctaacggtgtgaacgcgcattcaaaatgtgtctgcgatgtacctc	38552
Db	38551	GTGGCAATATGCTGCTGTACGCGTACCGCATTCMAAATGTTGTCTCGATGACCTTTC	600
Qy	601	ttgttgacattcacacacacagagcgtcatatcagcggctttaaagttgcgtgaacaggtg	38492
Db	38491	TTTGTGCGATTTCACCCACCAAGAGGTATATACAGCGGCTTAAACAGTGTGACAGGTGG	660
Qy	661	ttgggtlaaggtttggatttagcatcgtcacagcggagatcgtgcgctgtcgtggaacct	38432
Db	38431	TTGGTAAAGTGTGGATTTAGCATTCGTACAGGCGCATATGCTGCGCTGTGCGGATCTT	720
Qy	721	tgaatagccgcgcgccttgcattcttcgcgaacttctctgcgaacatcccccaagctc	38372
Db	38371	TGAATACCGCGCGCTTTGCTGATCTTCGACACTTCTTCGACAACTCTCCCAACAGCTC	780
Qy	781	ttgttgacattcacacacacagcgcgcgttaccatggaatctctgcactcttgcgcccg	38312
Db	38311	TGTTTGTGCAATATCAACCGCAGCGCTGTACACAGCAATCTGTGATCTTGGATCTTGG	840
Qy	841	cgctgcgcgaactacagcaatlaatcgcgaataagcaatgttgcgcacactgtgcactctt	38252
Db	38251	CGTCGCGGACATACCGCAATATATCGCATTAATCGCAATATGTTGCGAGCACTTGCACTT	900
Qy	901	tgccttagtatttcctttaaactgtcccc	38192
Db	38191	TGCTTACTATTTCTTCAACGCTTGGC	38164

RESULT 4

CVU39285/c

LOCUS

DEFINITION

CVU39285

Cloning vector TLF97-2, phage lambda lacZ translational fusion

ACCESSION

VERSION

U39285.1

KEYWORDS

GI:1066308

SOURCE

ORGANISM

Cloning vector TLF97-2.

artificial sequence; vectors.

REFERENCE

AUTHORS

1 (bases 1 to 42530)

St Pierre, R. and Linn, T.

TITLE

A refined vector system for the in vitro construction of

single-copy transcriptional or translational fusions to lacZ

96186904

2 (bases 1 to 42530)

Sclepierre, R.

JOURNAL

MEDLINE

Gene 169 (1), 65-68 (1996)

REFERENCE

AUTHORS

Submitted (24-Oct-1995) Thomas Linn, Microbiology and Immunology,

University of Western Ontario, London, Ontario N6A5C1, Canada

location/Qualifiers

1..42530

FEATURES

source

/organism="Cloning vector TLF97-2"

/db\_xref="taxon:43841"

terminator

complement(19404..19489)

gene

/note="rtnB2 terminator"

complement(19507..20760)

CDS

/gene="lacZ"

complement(19507..20760)

/codon\_start=-1

[illegible]





Query Match	Best Local Similarity	Matches	Score	DB 12:	Length	42704:
98.8%	Pred. No. 4.4e-283;	Mismatches	3;	Indels	0;	Gaps
925;	Conservative	0;				
1	aaagcgagaggtgtgtgtcgagacccctctgcacccgccttcacgaagtgatgta	60				
39266	AAACGCGAGAGGTTGTTAGCGCACCTCTGCACCCGCTTTCACGAAGTCAATGTGTA	120				
61	aaagcgagaggtgtgtgtcgagacccctctgcacccgccttcacgaagtgatgta	180				
39206	AAAGCGCGAGGCTAATTACTTAATGAATTCAGGACACAGTGGCTAACGGCTACAGTT	39147				
121	tgggtctgtctgtctgtcgagcgagacgaacgctcgtaacgatttgatgtaacggttcac	39087				
39146	TGGGTCTGTCTGTCTGTGGCGCGCATACCGCTGTACGCAATTTGGTGATTCGAGTTCTTC	240				
181	ttccggtatcgcattaatcaagacaacggaaagagacactgcttaacaaagctcgcgcac					

[illegible]

- TITLE Nucleotide sequence analysis of DNA. II. Complete nucleotide sequence of the cohesive ends of bacteriophage lambda DNA  
JOURNAL MEDLINE 71209066  
PUBMED 4931680  
REFERENCE 3 (bases 45493 to 45963)  
AUTHORS Imada, M. and Tsugita, A.  
TITLE Amino acid sequence of lambda phage endolysin  
JOURNAL MEDLINE 71209066  
PUBMED 4931680  
REFERENCE 4 (sites)  
AUTHORS Weigel, P.H., Englund, P.T., Murray, K. and Old, R.W.  
TITLE The 3'-terminal nucleotide sequences of bacteriophage lambda DNA  
JOURNAL MEDLINE 73215915  
PUBMED 4515613  
REFERENCE 5 (bases 38597 to 38672)  
AUTHORS Dahlberg, J.E. and Blattner, F.R.  
TITLE In vitro transcription products of lambda DNA: Nucleotide sequences and regulatory sites  
JOURNAL MEDLINE 1095210  
PUBMED 1095210  
REFERENCE 6 (bases 37945 to 38027)  
AUTHORS Maniatis, T., Plasmone, M., Backman, K., Kiehl, D., Flashman, S., Jeffrey, A. and Maurer, R.  
TITLE Recognition sequences of repressor and polymerase in the operators of bacteriophage lambda  
JOURNAL MEDLINE 75189495  
PUBMED 167018  
REFERENCE 7 (bases 35583 to 35600)  
AUTHORS Kiehl, D.G., Agarwal, K.L. and Khorana, H.G.  
TITLE The nucleotide sequence in the promoter region of the gene N in bacteriophage lambda  
JOURNAL MEDLINE 75189495  
PUBMED 167018  
REFERENCE 8 (bases 35434 to 35618)  
AUTHORS Dahlberg, J.E. and Blattner, F.R.  
TITLE Sequence of the promoter-operator proximal region of the major leftward RNA of bacteriophage lambda  
JOURNAL MEDLINE 76031664  
PUBMED 1178525  
REFERENCE 9 (bases 37945 to 38018)  
AUTHORS Maniatis, T., Jeffrey, A. and Kiehl, D.G.  
TITLE Nucleotide sequence of the rightward operator of phage lambda  
JOURNAL MEDLINE 75189495  
PUBMED 167018  
REFERENCE 10 (bases 44588 to 44773)  
AUTHORS Sklar, J., Yot, P. and Weissman, S.M.  
TITLE Determination of genes, restriction sites, and DNA sequences surrounding the 6S RNA template of bacteriophage lambda  
JOURNAL MEDLINE 75217747  
PUBMED 1098044  
REFERENCE 11 (bases 37905 to 37989)  
AUTHORS Malz, A., Pirrotta, V. and Ineichen, K.  
TITLE Lambda repressor regulates the switch between PR and P<sub>RM</sub> promoters  
JOURNAL MEDLINE 76267718  
PUBMED 958438  
REFERENCE 12 (bases 37946 to 38039)  
AUTHORS Smith, G.R., Eisen, R., Reichardt, L. and Hedgepeth, J.  
TITLE Rightward operator  
JOURNAL MEDLINE 76152323  
PUBMED 958438  
REFERENCE 13 (bases 35578 to 35667; 37903 to 38027)  
AUTHORS Plasmone, M., Backman, K., Humayun, M.Z., Jeffrey, A., Maurer, R., Meyer, B. and Sauer, R.T.  
TITLE Autoregulation and function of a repressor in bacteriophage lambda  
JOURNAL MEDLINE 76271154  
PUBMED 959843  
REFERENCE 14 (bases 35578 to 35667)  
AUTHORS Humayun, M.Z., Jeffrey, A. and Plasmone, M.  
TITLE Completed DNA sequences and organization of repressor-binding sites in the operators of phage lambda  
JOURNAL MEDLINE 77209970  
PUBMED 875019  
REFERENCE 15 (bases 38610 to 38732)  
AUTHORS Scherer, G., Hodom, G. and Kossel, H.  
TITLE DNA base sequence of the po promoter region of phage lambda  
JOURNAL MEDLINE 77100320  
PUBMED 834253  
REFERENCE 16 (bases 38041 to 38241)  
AUTHORS Roberts, T.M., Shimatake, H., Brady, C. and Rosenberg, M.  
TITLE Sequence of Cro gene of bacteriophage lambda  
JOURNAL MEDLINE 78071724  
PUBMED 593347  
REFERENCE 17 (bases 27616 to 28935)  
AUTHORS Davies, R.W., Schreier, P.H. and Buchel, D.E.  
TITLE Nucleotide sequence of the attachment site of coliphage lambda  
JOURNAL MEDLINE 78071724  
PUBMED 593399  
REFERENCE 18 (bases 37206 to 37263; 37914 to 37970)  
AUTHORS Humayun, M.Z.  
TITLE DNA sequence at the end of the CI gene in bacteriophage lambda  
JOURNAL MEDLINE 78011659  
PUBMED 909767  
REFERENCE 19 (bases 27617 to 27934)  
AUTHORS Landy, A. and Ross, W.  
TITLE Viral integration and excision: structure of the lambda att sites  
JOURNAL MEDLINE 77258934  
PUBMED 331474  
REFERENCE 20 (bases 39062 to 39170)  
AUTHORS Blattner, F.R.  
TITLE Physical structure of the replication origin of bacteriophage lambda  
JOURNAL MEDLINE 78054731  
PUBMED 929187  
REFERENCE 21 (bases 44467 to 44807)  
AUTHORS Sklar, J.L.  
TITLE Structure and function of two regions of DNA controlling the synthesis of prokaryotic RNAs  
JOURNAL MEDLINE 78234064  
PUBMED 354508  
REFERENCE 23 (bases 13 to 72; 48391 to 48502)  
AUTHORS Nichols, R.P. and Donelson, J.E.  
TITLE Lambda Nucleotide sequence surrounding the cos site of bacteriophage lambda  
JOURNAL MEDLINE 78197067  
PUBMED 668898  
REFERENCE 24 (bases 37938 to 38016; 35589 to 35666)  
AUTHORS Flashman, S.M.



TITLE Mutational analysis of the operators of bacteriophage lambda  
JOURNAL Molecular & general genetics : MGG. 166 (1), 61-73 (1978)  
MEDLINE 79114073  
PUBMED 368570  
REFERENCE 25 (bases 37990 to 38982)  
AUTHORS Schwarz, E., Scherer, G., Hobom, G. and Kossel, H.  
TITLE Nucleotide sequence of cro, cII and part of the O gene in phage lambda DNA  
JOURNAL Nature. 272 (5652), 410-414 (1978)  
MEDLINE 78135462  
PUBMED 264238  
REFERENCE 26 (bases 38212 to 38362)  
AUTHORS Rosenberg, M., Court, D., Shimatake, H., Brady, C. and Wulff, D.L.  
TITLE The relationship between function and DNA sequence in an intergenic regulatory region in phage lambda  
JOURNAL Nature. 272 (5652), 414-423 (1978)  
MEDLINE 78135463  
PUBMED 634366  
REFERENCE 27 (bases 37224 to 37940)  
AUTHORS Sauer, R.T.  
TITLE DNA sequence of the bacteriophage gamma cI gene  
JOURNAL Nature. 276 (5685), 301-302 (1978)  
MEDLINE 79053284  
PUBMED 714163  
REFERENCE 28 (bases 38597 to 39688)  
AUTHORS Scherer, G.  
TITLE Nucleotide sequence of the O gene and of the origin of replication in bacteriophage lambda DNA  
JOURNAL Nucleic acids research. 5 (9), 3141-3156 (1978)  
MEDLINE 79033241  
PUBMED 704348  
REFERENCE 29 (bases 29711 to 29811; 31043 to 31058)  
AUTHORS Davies, R.W., Schreier, P.H. and Buchel, D.E.  
TITLE Determination of the endpoints of partial deletion mutants of the attachment site of bacteriophage lambda by DNA sequencing

Query Match 98.8%; Score 923.2; DB 7; Length 48502;  
Best Local Similarity 99.7%; Pred. No. 4.4e-283;  
Matches 925; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 aaagcgagagaggtttagcgagcactctgacacccgcttcaagaagtcagtgtga 60  
45064 AAAGCGAGAGAGGTTTAGCGGCGACCTCCGCCCGCTTTCACGAAGTCATGTGTA 45005  
QY 61 aaagcgagagaggtttagcgagcactctgacacccgcttcaagaagtcagtgtga 120  
45004 AAAGCGAGAGAGGTTTAGCGGCGACCTCCGCCCGCTTTCACGAAGTCATGTGTA 44945  
QY 121 tgggttgctgtctgctgagcgagatgaagcgcctgaagcatttggtgacgcgttctgc 180  
44944 TGGGTGTGCTGTGCTGCGCGCGATGACCGCTGTAGCATTTGGTGTGCTGCTGCG 44885  
QY 181 ttcggatctgcttaattcagacacgaaggaagacacgtgcttaaccagtcgcgcgac 240  
44884 TTCGGATCTCCTTAATTCAACACAGGAAGACACTGGCTACCGCTGCGGAC 44825  
QY 241 tcttcagatttgacatgaatgccttcttcttctgctgacatataaaatccgaagac 300  
44824 TCTTCAGATTGACATGAATGCTCTTACCTGTTGTGACAGATATAAAATCCGGAAC 44765  
QY 301 cgtatgagagctctactactactacacgcgaactgttcggagatgcatttgcagactc 360  
44764 CGTATGAGAGCTCTACTACTACTACACGCAGACTGTTCGGAGATGTCATTTGCAACCTC 44705  
QY 361 tctgctcgatgagtggtgaggttccagacagatcgtcgaagtgaaccaactgaagcgaa 420  
44704 TCTGCTCGATGAGTGGTGAAGTTCAGACGATACGTCGAAGTACCACTAGCGGAATCG 44645  
QY 421 gtagtaagcgccgctcttctcactacacacgaagcgaaatlaaccatcgttga 480  
44644 GTAGTAAGCGCCGCTCTTCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 44585  
QY 481 gtcaaatlaaccatlaattcaataaagtaacatcgcgttcaataatgttgcatcc 540

Db 44584 GTCAAAATTAACCAATTTTATTCATTAAGTCAATATCAATGCGTTAATATGTCATCC 44525  
QY 541 gtgcaatcatgctgctgaacgctgacacgcatcacaatgtgtctgcatgctctc 600  
44524 GTGCAATCATGCTGCTGAACGCTGACACGCGATCAAAATGTTGCTGGATGACTCTTC 44465  
QY 601 tctgtagcatgacacacacgagcgtcatagacgcttcaacagcgtgacgagtgag 660  
44464 TCTGTGCTAGTTCACACACGAGCGCTCATACAGCGGCTTAACAGTGCCTGACAGCTG 44405  
QY 661 ttggaaggttgaggatgaacatcgttcacagcgagatgctgctgctgcatctc 720  
44404 TTGGAAGGTGAGGATGAACATCGTTCACAGCGGATATGCTGCGCTTGGCATCTC 44345  
QY 721 tgaatagcagcgccttgcattcttcgcacacttctcgaacatcccccacagctc 780  
44344 TGAATAGCAGCGCCTTGCATCTTCGACACTCTCCACAGCTC 44285  
QY 781 ttttggcaataatcaacgcagcgcctgtacatgacatctcgtacatctgacatctg 840  
44284 TTTTGGCAATATCAACCGCAGCGCTGTACATGACATCTTCGATCTGCCCCGG 44225  
QY 841 cgtcgagcactacgagcaataatccgataagcgaatgttgagagcatttgacactt 900  
44224 CGTCGAGCACTACGCAATATCCGATTAAGCAATGTTGCGAGCACTTGAGTACTT 44165  
QY 901 tgccttgattcttcctcaacgtgcgc 928  
44164 TGCTTGTATTTCTTCAAGCTTGGC 44137

RESULT 8  
LOCUS U02453 9170 bp DNA SYN 29-MAR-1996  
DEFINITION Cloning vector lambda EMBL3, right arm.  
ACCESSION U02453  
VERSION U02453.1 GI:413819  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE 1 (bases 1 to 9170)  
AUTHORS Kitz, P.A.  
TITLE Cloning vector lambda EMBL3.  
JOURNAL Cloning vector lambda EMBL3.  
REFERENCE 2 (bases 1 to 9170)  
AUTHORS Frischauf, A.M., Lehnach, H., Poustka, A. and Murray, N.  
TITLE Lambda replacement vectors carrying polylinker sequences  
JOURNAL J. Mol. Biol. 170 (4), 827-842 (1983)  
MEDLINE 84064856  
REFERENCE 3 (bases 1 to 9170)  
AUTHORS Kitz, P.A.  
TITLE Direct Submission  
JOURNAL Submitted (07-OCT-1993) Paul A. Kitz, Clontech Laboratories, Inc., 1020 East Meadow Circle, Palo Alto, CA 94303, USA  
COMMENT This sequence has been compiled from information in the sequence databases, published literature and other sources. If you suspect there is an error in this sequence, please contact Clontech's Technical Service Department at (415) 424-8222 or (800) 662-2566, extension 3 or E-mail TECH@CLONTECH.COM.  
FEATURES  
source location/Qualifiers  
1. 9170  
BASE COUNT 2534 a 1996 c 2251 g 2389 t  
ORIGIN  
/organism="Cloning vector lambda EMBL3"  
/db\_xref="taxon:31783"

Query Match 98.7%; Score 921.6; DB 12; Length 9170;  
Best Local Similarity 99.6%; Pred. No. 1.4e-282;  
Matches 924; Conservative 0; Mismatches 4; Indels 0; Gaps 0;





```

|||||
Db 691 TCATGCCGTTAATATG--TGCATCCGTCGCGATCATGCGTAAAGTGTGATCCATTCA 633
Qy 576 aaatgtgtcgtcgtatgactcttctgtgacattgacacacagagctcacaagcg 635
Db 632 AAATGTTGTCGTGGATTGACTCTTCTTGTGTCGATTCGACACAGAGCTCATACAGCG 573
Qy 636 gcttaacagtgctgaacagagtgagtgtaagatttgagattagatcgtacacagcg 695
Db 572 GCTTAACAGTGGCTGACACAGTGGGTGGTAAGTTGGGATTAAGTACGTCACAGCGC 513
Qy 696 gatactgtcgtcgtcgtgacatcccttaataagcagacgtcttgacatcttcgacatct 755
Db 512 GATATGCTGCGCTGCTGCGATCCTTGAATAGCCGAGCGCTTGCATCTTCGCGACTCTT 453
Qy 756 tctcgacaactctcccccaagctctgttttgacaatacaccgacgacctatccat 815
Db 452 TCTGCACAACCTCTCCCGACAGCTGCTGTTTGGCATATACACCCGACGCTGTACCAT 393
Qy 816 ggcacatcttgcacatcttgcaccccgctgcgacgacgacgacgacgacgacgacgacg 875
Db 392 GGCATCTCTGCATCTTGCCTCCCGGCTGCGCCACTACGCGCAATATCCGATTAAGCA 333
Qy 876 atgttcgagacattgcagtaacttgccttagtaattcctcctaagctgcgc 928
Db 332 ATGTTGCGACACTTCGACGTAAGTCTTGGCTTGTCTTCTTCAAGCTTGC 280

RESULT 12
SIHK0220 805 bp DNA PHG 06-OCT-1992
LOCUS Colliphage HK022 0 gene with late antitermination function.
DEFINITION X60309.1 GI:15751
ACCESSION X60309.1 GI:15751
VERSION X60309.1 GI:15751
KEYWORDS antitermination block, Q gene; rpoB60; rpoB60 mutation.
SOURCE Enterobacteria phage HK022.
ORGANISM Enterobacteria phage HK022.
REFERENCE 1 (bases 1 to 805)
AUTHORS Atkinson, B.L.
TITLE Lambda phage group.
JOURNAL Submitted (05-SEP-1991) B.L. Atkinson, Columbia Univ College of Physicians and Surgeons, 701 West 168th Street, New York NY 10032, USA
2 (bases 1 to 805)
REFERENCE Atkinson, B.L. and Gotesman, M.E.
TITLE The Escherichia coli rpoB60 mutation blocks antitermination by coliphage HK022 Q-function
JOURNAL J Mol. Biol. 227 (1), 29-37 (1992)
9239369
FEATURES
source
Location/Qualifiers
1..805
/organism="Enterobacteria phage HK022"
/db_xref="taxon:10742"
34..657
/gene="Q"
34..657
/gene="Q"
34..657
/codon_start=1
/transl_table=1
/protein_id="CAA43854.1"
/db_xref="GI:15752"
/db_xref="SPTREBL:002582"
/translation="MARLESYAKFHSPKSPMSDSPPTASDLSSTDYMAAMGAOSQ
AGFGMAACGAGHLSNDKQKAINYLMOFAHVKYKQVAKLENTKAKVLQVLAIF
YRAVTMLIPNLTQPTSRVTKPLYLALVYQCHKEKSIADNILNAVAR"
misc_feature
700..729
/feature="PR_35 and -10"
prim_transcript
736..792
/standard_name="pr"
stem_loop
766..788

```

```

BASE COUNT 229 a 185 c 216 g 175 t
ORIGIN
Query Match
Best Local Similarity 95.18; Score 436.2; DB 7; Length 805;
Matches 450; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
Qy 456 aacggcgaaatcaaccacgcttgatcgaattaccacattatcaataagcaata 515
Db 757 AACGAGTCATATTATTCATGCTGGGCAATATACCAACTTATTCATAAAGTCATA 698
Qy 516 tcatgcgttaataatgttgcacatccgttggaacatcagctgtcaacgctgtaacgacatca 575
Db 697 TCATGCCGTTAATATGTTGTCGATCCGTCGCAATCAATGCTTACGTCGACCGCATTC 638
Qy 576 aaatgtgtcgtcgtatgactcttctgtgacattgacacacagagctcacaagcg 635
Db 632 AAATGTTGTCGTGGATTGACTCTTCTTGTGTCGATTCGACACAGAGCTCATACAGCG 573
Qy 636 gcttaacagtgctgaacagagtgagtgtaagatttgagattagatcgtacacagcg 695
Db 572 GCTTAACAGTGGCTGACACAGTGGGTGGTAAGTTGGGATTAAGTACGTCACAGCGC 513
Qy 696 gatactgtcgtcgtcgtgacatcccttaataagcagacgtcttgacatcttcgacatct 755
Db 512 GATATGCTGCGCTGCTGCGATCCTTGAATAGCCGAGCGCTTGCATCTTCGCGACTCTT 453
Qy 756 tctcgacaactctcccccaagctctgttttgacaatacaccgacgacctatccat 815
Db 452 TCTGCACAACCTCTCCCGACAGCTGCTGTTTGGCATATACACCCGACGCTGTACCAT 393
Qy 816 ggcacatcttgcacatcttgcaccccgctgcgacgacgacgacgacgacgacgacgacg 875
Db 392 GGCATCTCTGCATCTTGCCTCCCGGCTGCGCCACTACGCGCAATATCCGATTAAGCA 333
Qy 876 atgttcgagacattgcagtaacttgccttagtaattcctcctaagctgcgc 928
Db 332 ATGTTGCGACACTTCGACGTAAGTCTTGGCTTGTCTTCTTCAAGCTTGC 280

RESULT 13
AF069308/4 40751 bp DNA PHG 17-OCT-2000
LOCUS Bacteriophage HK022, complete genome.
DEFINITION AF069308
ACCESSION AF069308
VERSION AF069308.1 GI:6863111
KEYWORDS
SOURCE
ORGANISM Enterobacteria phage HK022.
REFERENCE 1 (bases 1 to 40751)
AUTHORS Hendrix, R.W., Ford, M.E., Duda, R.L., Youlton, A., Hatfull, G.F. and
TITLE Genomic sequences of bacteriophages HK97 and HK022: pervasive
JOURNAL J Mol. Biol. 299 (1), 27-51 (2000)
2038598
REFERENCE 2 (bases 1 to 40751)
AUTHORS Hendrix, R.W., Ford, M.E., Duda, R.L., Youlton, A., Hatfull, G.F. and
TITLE Direct Submission
JOURNAL Submitted (31-MAY-1998) Pittsburgh Bacteriophage Institute & Dept.
15260, USA
of Biological Sciences, University of Pittsburgh, Pittsburgh, PA
FEATURES
source
Location/Qualifiers
1..40751
/organism="Enterobacteria phage HK022"
/db_xref="taxon:10742"

```



```

CDS             ERV1
                8478..10898
                /note="gp16"
                /codon_start=1

```

OY	456	aacgacggaatlaaccacacgcttgagtaacaaatlacccaatttataatgaatgaata	515
Db	36934	AACAGAGTCATTATATTCATCGTTGGGTCAAATTTACCAACTTATTCAAAAAAGTCAATA	36875
OY	516	tcatgcgcttaataatgttgcacacgcgttggaatacatgctgcctaaagttgacgcgattca	575
Db	36874	TCATGCCGTTAAATATGTGTCCATCCGTGGCAATCATAGCTGCTACGTGGAGCGGATTCA	36815
OY	576	aaatggtctcgtgattgaattcttcttcttgatcatgacacacagagcgtatcaacg	635
Db	36814	AAATGTTGTCTGCCATTGACTCTCTCTTGTGCATATGCACCCAGACGTCATATACAGC	36755
OY	636	gcttaacagtcgctgacacagtggtggttggtaagtttggattagcaatgcgtcaacgac	695
Db	36754	GCTTAAACATGCGTGACCCAGGTGGGTGGATTAAGCTTTGGATTAAGCATCTTTACACGC	36695
OY	696	gatatcgtcgtcttgcttgatcactcttgaaatagacgaagccttgcattcttcgaactct	755
Db	36694	GATATCGCGCGCTTCTGCACTCTCGAATATAGCCGACCCCTTTCATCTTCCGACACTT	36635
OY	756	tctgcacaactctcccccacagctctgttttggcaatalaaccgacggcctgttaecat	815
Db	36634	TCCTGACAACTTCCCCACACGCTCTGTTTGGCAATATACAGCGCCCGGACAGTACCGT	36575
OY	816	ggcaatctctgcacatcttgcctccctggcgctcgcgacactaagcgaaataatccgaataagca	875
Db	36574	GGCATATCTCTGCANCTTGCCGCCCGGTCTCGCGACATACGGCAATATATCCGATAGGCA	36515
OY	876	atgttgcgaagacttgcaatgaacttgccttgaattatcttcctaagcttgccc	928
Db	36514	ATGTTGGAGAACTTGCACATCACTTGGCTTACTATTTCTCTAACCTTTGGC	36462

viruses; dsDNA viruses, no RNA stage; Tailed phages; Siphoviridae.

terminator	complement(join(1..23,38289..38297))
gene	/note="terminator for prophage excision transcript: XIS complement(50..1207)
	/gene="int"
	/gene="hkaA"
CDS	/note="hkaA complement(50..1207)

feature	start	end	score	strand	gene	transcript	transcript_start	transcript_end	transcript_score	transcript_strand	transcript_gene	transcript_transcript
repeat_region	1207	1232		+								
promoter	1207	1232		+								
repeat_region	1233	1253		+								
gene	1233	1253		+								
CDS	1233	1253		+								

```
misc_feature      1368. .1383
repeat_region     1441. .1463
```

```

promoter
  complement(1664..1702)
  /function="phage integration"
protein_bind
  complement(1689..1701)
  /bound_moiety="cII protein"

```



Db	15317	TCTGCAGACATCTACCCACACACATTCGTTTGGCATATACAGGGCCCCGGCCAGTACCGT	15258
OY	816	ggcaatctcgcgacatcttggcccccgcgcttcgcgcgacacgcgcaataatccgataagcga	875
Db	15257	GACATCTCTGCATTTGGCGCCGGCGCTCGCGGACCTACGGCATATATTCGGCTTAAGCA	15198
OY	876	atgtctggagcaacttcgactactcttgcttaagatattcttccttaagctgcgcc	928
Db	15197	ATGTTGGGACACTTCACACACTTTCGCTTAGTATTTTCCTTCGAGCTTTGGC	15145

RESULT	15	
AF069529/c		
LOCUS	AF069529	39732 bp
DEFINITION	Bacteriophage HK97, complete genome.	PHG
ACCESSION	AF069529	18-OCT-2000
VERSION	AF069529.1	
KEYWORDS	GI:5901584	
SOURCE		
ORGANISM	Bacteriophage HK97.	
	Bacteriophage HK97	

REFERENCE  
1 (bases 1 to 39732)

**TITLE** Genomic sequences of bacteriophages HK97 and HK022: pervasive genetic mosaicism in the lambdoid bacteriophages  
**JOURNAL** J. Mol. Biol. 299 (1), 27-51 (2000)  
**PMID** 20328508  
**MEDLINE**

TITLE  
 Journal  
 Submitted (01-JUN-1998) Pittsburgh Bacteriophage Institute &  
 Direct Submission  
 Hendrix, R.W., Foulds, M.E., Duda, R.L., Younton, A., Hatfull, G.F. and

Department of Biological Sciences, University of Pittsburgh, Pittsburgh, PA 15260, USA

FEATURES	Location/Qualifiers
source	1. .39732

```

/organism="Bacteriophage HK97"
/db_xref="taxon:37554"

```

CDS /gene1 50.535

```

/ncore="gpl"
/codon_start=1
/transl_table=1
/evidence=not_experimental
/product="terminase small subunit"
/protein_id="AAR31097.1"
/db_xref="GI:6901594"
/translation="MADKRIKRSDSAAAVOAMKNAADVITDPPSHAGLEKKAEPWPHI
NISKALDSWTLPADLLAAVLANNOLITVLRDLRKEKIRIGEDEDELIDLKQ
VELQRTLLAQRDQLQHSATNGESRDQKKRQNDKARDARTKNEHDDDDNLIAPPKI
G"
gene 542..2056
/feature="2"
542..2056
CDS

```

```

gene
cds

542..2056
/gene="2"
542..2056
/gene="2"
/note="gp2"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="terminase large subunit"
/protein_id="AF31098.1"
/db_xref="GI:6901595"
/translacion="MTRGRVATIERECIYPEEGKLTQPMRLDPFOKDTLLAVDNP
ACIDAIILSIARKNCTGLIAGILLAHVGPAYVNOTIVSGALSREQALVPMIDAWK
MYNLNKLOEIVHITPSCGKRLIGLPCNVEYALSAEGKTYGLSPILAIIETGQVGG
PDDIDTALITAOGAENPLIVISTOAMADLLSIWIDYAVRSKDPIVCHYEAP
KADISIKRESMLANDPALTGEFERSEKDMAROEKGRMSPFENFRINLTNRVSYSE
FLTSRWELGCEMPINTPRKMYAGLDLSARNDLVALYAGADGCVMDVFEFFWTPK

```

gene  
CDS

terminator

gene	3348. .4025
	/gene="4"
CDS	3348. .4025

gene  
CDS

terminator  
gene  
CDS

TLEERTTDRAPYDVWVREGILLRTTPGASVDYSEVVAADIAEITIGDFDITSMAFDRWR  
 DOFRKADADILSLPIVEFOGFKMGPAVDLTESLIMGRRVRGHMPVLTMCVAVNA  
 VYKDAGNRKIDLKSKATGRIDGMVAMTMSVGAANGEVTEGGGDFDDFIFRPLSM"  
 2056. .3330

/gene="3"  
 /note="Gp3"  
 /codon\_start=1  
 /transl\_table=1  
 /evidence=experimental  
 /product="portal protein"  
 /protein\_id="AA03108.1"  
 /db\_xref="GI:6901585"  
 /translation="MEEPTPTIDRTNNGMARLKSMEFVGRLVTPNPGSGTGPVSAAH  
 GIADDSIDIERLLITSTYWRCSLLSTLTACPLDIVEETPDNDNRKRVLDNSPLAR  
 LRSNPQVIAQDFEARMQMOLCFYGNAAVAIDNRNSAGDVSTLPILOANNDVLRK  
 KVVYRQVROISEYDEFQSKELFHLKGFSGFLGSLSPINAPACKSAGVAANAMDQDRDF  
 ANAKSPQILSTEKLTAEFGSGSYDNEBNKEITAGCPYKRLRLILEAGTSTSAITVTPQ  
 DAMKASRKFOVSEKLTAEFGSGSYDNEBNKEITAGCPYKRLRLILEAGTSTSAITVTPQ  
 ENSIQWMLIPSKRYGVGLAEHNLIDGLGKDSKSRRAAFKAGSGELRTINERTPTDMMN  
 PLPGSGDVAARQIQVPTIDTGTNKEPRNNGA"  
 3309.

```

/ano="4"
3348. .4025
/gene="4"
3348. .4025
/ano="4"

```

```

/function="head maturation"
/note="Sp4"
/codon_start=1
/transl_table=11
/evidence=experimental
/product="protease"
/protein_id="AAE1089.1"
/db_xref="GI:6901586"
/transl_table="mpriyutlsfpdetrklfgtggkkgcigfegyaafvnttdsdgltitl
pgeafamgfnalantnrtkrvaefnfhmtkmlrpygkksdladpbrklylrgolrpgshgaidk
paaamgnfyalnecsgvsaanddtytlptgtrikrnlaolrealtvtpfpannqoaiams
pmsnifetrvdenmlrbsvgltlrsqavgliafrkksairseesbgdgnaeoinmllosikms
psnsgkr"
/size .5185
/length 4028 .5185

```

```

/transl_table=11
/evidence=experimental
/product="major head subunit precursor"
/protein_id="AAFA31090.1"
/db_xref="GI:6901587"
/translation="MSLALILQKAIIESQOKMTQLPDAQKAEISTGCVSKLOSLDLK
KVEELISGRFLDEKCKLASGANGENGEKSPSEAAEELISMDGCKGFFGKRTFN
KSLSDASASLLOPMQIPGIIIMPGLRLLIDLLAQRSTSNLLEVRVEPTNNNA
DVVAKARAKPSDITFSKQRTANVTITAHVOASROYMDAPMLQSYINRRMLTAKL
EGGOLLNDGGDNGDNEGINVYANATYDPSLNAQSDPRADIIAIIAYQTESFSNGIV
LNPFRMHNILATLKIKNEGYIEGGPQAFSTNMGFLPYVPTKQAQAGITVGGCFPMASO
VMDRRDAIVESRDRDNFVKNNMLTLCERLALAHRTPLATIKGTFSSGS"
5192..5213
/feature="putative"
5219..5545
/gene="6"
5219..5545
/gene="6"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="Gp6"
/protein_id="AAFA31099.1"
/db_xref="GI:6901596"
/translation="MADIVLVIVSLFKQOIEFEDDRDLITVQAQAFDYCKMC
DEPMKVAADIIPAAYGAVLVLPADMFELTRAQSSQVLYENAAERMMFLTRHWKRA

```



gene  
 5545..5883  
 /gene="7"  
 5545..5883  
 /gene="7"  
 /note="Gp7"  
 /codon\_start=1  
 /transl\_table=11  
 /evidence=not\_experimental  
 /product="putative head-tail adaptor"  
 /protein\_id="AA01597"  
 /db\_xref="GI:6901597"  
 /translation="MEGRFRNRVKILFPTTSRDPSGQPVESMTGNGNPVPAEVGIG  
 REOLSGAETAAQATIRVMMRFRSELNASSRLVSGPYKGVOLNIGPVPANATGTRL  
 EILCTGAER"  
 5880..6329  
 /gene="10"  
 /gene="6329  
 5880..6329  
 /gene="10"  
 /codon\_start=1  
 /transl\_table=11  
 /evidence=not\_experimental  
 /product="Gp10"  
 /protein\_id="AA01101.1"  
 /db\_xref="GI:6901598"  
 /translation="MIFTSIDFGLNDIAKDLEALSRANNNKYLDRDPAAGAEVLKEE  
 VIDRAPVETGLKNNVYVYVOKSRREGTISGVHIGVNPRTGNSDNTKKANNPRNAF  
 YMRVELGTANMPAHFVRPAIDTREELASVALARNQAIDVLSK"  
 6326..6673  
 /gene="11"  
 6326..6673  
 /gene="11"  
 /codon\_start=1  
 /transl\_table=11  
 /evidence=not\_experimental  
 /product="Gp11"  
 /protein\_id="AA031102.1"  
 /db\_xref="GI:6901599"  
 /translation="MNEDNIYALSLPLAEGRYVYVAPLGSQKPSVSPMIIIFSIVD  
 DVSADVLGQAESRVSVQVDYVSTSIASRLRLDLVLADEPLTPEVVKIPGEPEY  
 RLYRATIDFKVTP"  
 6692..6719  
 /note="putative"  
 /evidence=not\_experimental  
 6733..7437  
 /gene="12"  
 6733..7437  
 /gene="12"  
 /note="Gp2"  
 /codon\_start=1  
 /transl\_table=11  
 /evidence=experimental  
 /product="major tail subunit"  
 /protein\_id="AA031091.1"  
 /db\_xref="GI:6901588"  
 /translation="MSALYKSQLTKILISLPSTKETMDSATFLDLSCTKEIOFTG  
 GOKODIVTFLCSTEOENINGSLSPSEISLGNFYNNPQDALRDAYNDITTYGFOIT  
 FPSGNGRFLAEVROHMSSTGNGVAAATSLRLKGPVPIIDSVLKLITDLPSSLVA  
 VGAISMVAVVAAAGKRPYATYTKKASTVSGQTSDFNKATAVSGDAGDYTCVTDSS  
 SPKRTVTSAACTTGIS"  
 7443..7462  
 /note="putative"  
 7472..7858  
 /gene="13"  
 7472..8145  
 /gene="14"  
 7472..7858  
 /gene="14"

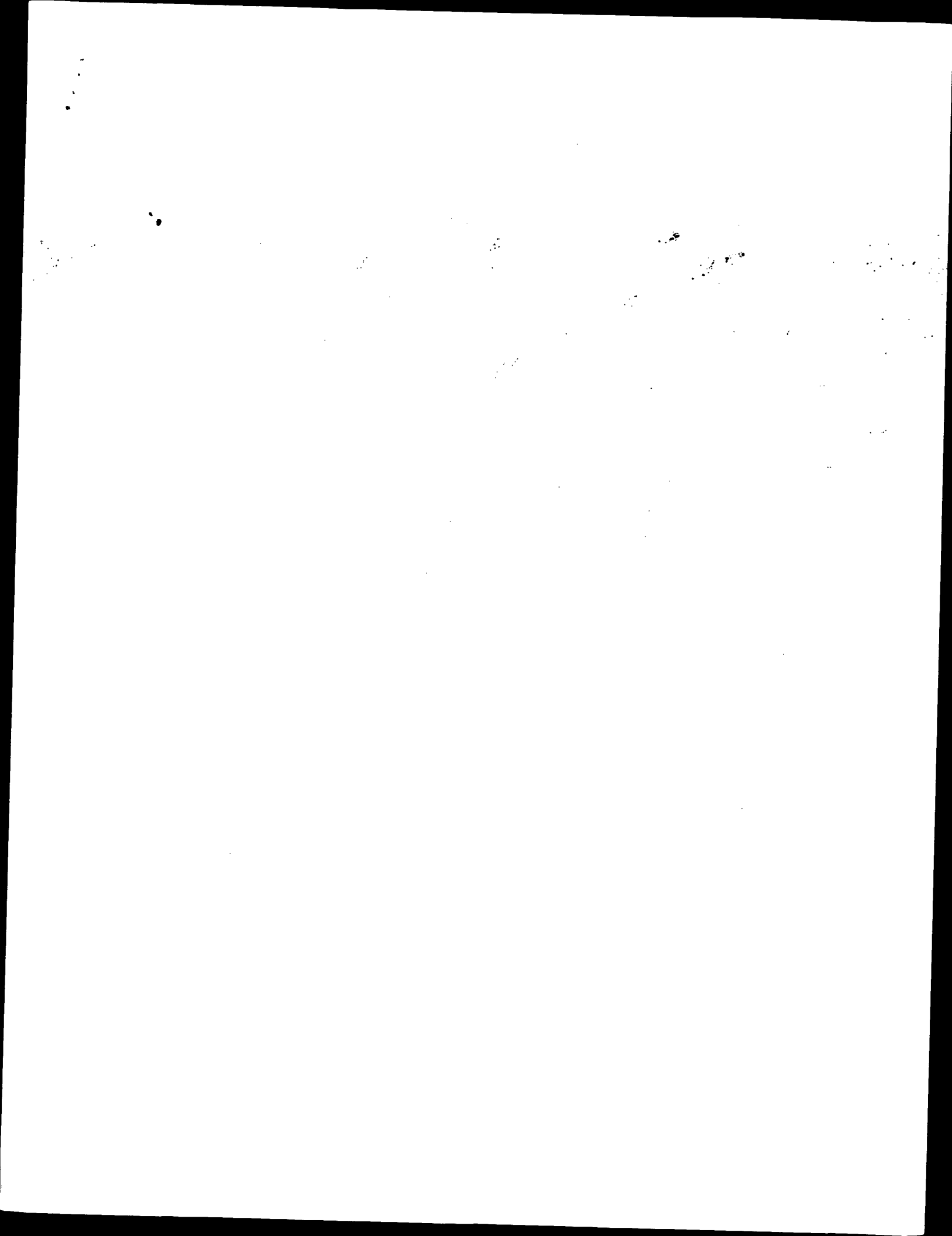
Query Match 45.7%; Score 426.8; DB 7; Length 39732;  
 Best Local Similarity 95.1%; Pred. No. 1.2e-124; Indels 1; Gaps 1;

Matches 451; Conservative 0; Mismatches 22;

QY 456 aacgagcgaattacccatcgttga-gtcaaatattacccaatttattcacaataagtcatt 514

Db 35920 AACGAGTGCATTTATCCATGTTGATGTCAAATTTACCAACTTATTCACAAAGTCAAT 35861  
 QY 515 atcagcgtlaatalatgttgcacatccgtggaatacatatgctgtaacgtgttgcacatc 574  
 Db 35860 ATCATGCCGTTATATGTTGCCATCCGTGGCAATCATGCTGCTAACGtgtgACCGCTTC 35801  
 QY 575 aaatgttgcgtggaattgactctcttgttgcatgtgacacacagagcgtlcatcacg 634  
 Db 35800 AAATGTTGTCCTCAATGACTCTTCTTGTGCTTGTGCAATTGACACACAGACGTCATPACAGC 35741  
 QY 635 ggcctaacgtgctgacacagctgagctgagctgagctgagctgagctgagctgagctgagc 694  
 Db 35740 GGCTTACAGTGCCTGACACAGGTGGGTTGGTGAGGTTGGATTAGCATGCTTACAGCG 35681  
 QY 695 cgaatgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 754  
 Db 35680 CGATATGCGGCGCTTGGCTGGCATTTGGAATAGCGACACCTTTGCATCTCCGCAATCT 35621  
 QY 755 ttctcgacaactctcccccacagctctgttlltgcaatalatcaacgacgagcgttgcac 814  
 Db 35620 TTCTCAACACTCTCCCCACAGCTCTGTTTGGCAATATCAACCGCGCTGTACCG 35561  
 QY 815 tggcaatctctgacatcttgcctcccggtgctgagcactacaggaataatcgcataagc 874  
 Db 35560 TGGCAATCTCTGCAATCTTGCCTCCGCGCTCGGCGACTACGCAATATATCGCATTAAGCG 35501  
 QY 875 aatgttgcgagcactgcaacttgccttgccttgccttgccttgccttgccttgccttgc 928  
 Db 35500 AATGTTGCGAGCACTTGCAGTACCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCTTGC 35447

Search completed: January 8, 2002, 17:22:58  
 Job time: 12469 sec



Run on: January 9, 2007  
(without alignments)  
1867.271 Million cell updates/sec

Title: US-09-740-211-13\_COPY\_11000\_11933  
 Effect score: 934  
 ctccaacgtgccccctgcagg 934

Scoring table:

	IDENTITY_NUC	%
Gapop 10.0		1.0
Gapext 1.0		1.0

Searched: 930621 seqs, 428002012 records  
 Translators: 1861242

Total number of hits satisfying chosen parameters:

```

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

```

post-processing:	Minimum	Match	0%
	Maximum	Match	100%

Maximum value: 45 summaries  
Listing first 45 summaries

```
Database : N_Geneseq_1101:*
1. /SIP52/qcdata/geneseq/geneseqn/NA1980.DAT:*
1. /SIP52/qcdata/geneseq/geneseqn/NA1981.DAT:*
```

- 1: /SIDS2/gcgdata/genseq/genseq/NA1980.DAT \*
- 2: /SIDS2/gcgdata/genseq/genseq/NA1981.DAT \*
- 3: /SIDS2/gcgdata/genseq/genseq/NA1982.DAT \*
- 4: /SIDS2/gcgdata/genseq/genseq/NA1983.DAT \*
- 5: /SIDS2/gcgdata/genseq/genseq/NA1984.DAT \*
- 6: /SIDS2/gcgdata/genseq/genseq/NA1985.DAT \*
- 7: /SIDS2/gcgdata/genseq/genseq/NA1987.DAT \*
- 8: /SIDS2/gcgdata/genseq/genseq/NA1988.DAT \*
- 9: /SIDS2/gcgdata/genseq/genseq/NA1989.DAT \*
- 10: /SIDS2/gcgdata/genseq/genseq/NA1990.DAT \*
- 11: /SIDS2/gcgdata/genseq/genseq/NA1991.DAT \*
- 12: /SIDS2/gcgdata/genseq/genseq/NA1992.DAT \*
- 13: /SIDS2/gcgdata/genseq/genseq/NA1993.DAT \*
- 14: /SIDS2/gcgdata/genseq/genseq/NA1994.DAT \*
- 15: /SIDS2/gcgdata/genseq/genseq/NA1996.DAT \*
- 16: /SIDS2/gcgdata/genseq/genseq/NA1997.DAT \*
- 17: /SIDS2/gcgdata/genseq/genseq/NA1998.DAT \*
- 18: /SIDS2/gcgdata/genseq/genseq/NA1999.DAT \*
- 19: /SIDS2/gcgdata/genseq/genseq/NA2000.DAT \*
- 20: /SIDS2/gcgdata/genseq/genseq/NA2001.DAT \*
- 21: /SIDS2/gcgdata/genseq/genseq/NA2001.DAT \*
- 22: /SIDS2/gcgdata/genseq/genseq/NA2001.DAT \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	100.0	11933	21	AAD001121	Recombinant adeno
2	934	11937	21	AAD086112	Human factor VIII
3	36.6	3.9	3785C	AAAI11992	S. cellulosum DNA
4	33	3.5	1063	AAAC37411	Arabidopsis thaila
5	31.32	3.4	49617	AAAF28541	Genomic fragment #
6	31.32	3.4	469	AAI12818	Probe #2751 for ge
7	31.38	3.4	469	AAI34173	Probe #2859 used t
8	31.8	3.4	469	AAI02735	Probe #2726 used t
9	31.8	3.4	542	AAI18836	Probe #8769 for ge
10	31.8	3.4	542	AAI14358	Probe #12644 used
11	31.8	3.4	936	AAAF58252	Oligonucleotide D

C	12	31.8	3.4	936	22	AAE58254
C	13	31.8	3.4	936	22	AAE58257
C	14	31.8	3.4	936	22	AAE58259
C	15	31.8	3.4	936	22	AAE58262
C	16	31.8	3.4	938	22	AAE58255
C	17	31.8	3.4	1698	19	AAE50577
C	18	31.8	3.4	4455	18	AAE255095
C	19	31.8	3.4	8709	22	AAE25628
C	20	31.6	3.4	247	21	AAC24010
C	21	31.6	3.4	4668	19	AAE60918
C	22	31.4	3.4	936	22	AAE58252
C	23	31.4	3.4	936	22	AAE58257
C	24	31.4	3.4	936	22	AAE58257
C	25	31.4	3.4	936	22	AAE58259
C	26	31.4	3.4	936	22	AAE58255
C	27	31.4	3.4	938	22	AAE58255
C	28	31.2	3.3	2940	18	AAT90501
C	29	31.2	3.3	3691	18	AAE905300
C	30	31.2	3.3	4147	18	AAT668934
C	31	31.2	3.3	4870	18	AAE606041
C	32	30.8	3.3	642	22	AAE71385
C	33	30.8	3.3	822	22	AAE71385
C	34	30.8	3.3	1951	16	AAE94449
C	35	30.8	3.3	1951	18	AAE49507
C	36	30.8	3.3	1951	19	AAE21443
C	37	30.8	3.3	1951	21	AAE289171
C	38	30.8	3.3	1951	21	AAE291005
C	39	30.8	3.3	6526	21	AAE249723
C	40	30.8	3.3	349880	22	AAE68527
C	41	30.6	3.3	3875	12	AAO11561
C	42	30.6	3.3	10732	21	AAAI0594
C	43	30.4	3.3	1130	21	AAE39064
C	44	30	3.2	698	21	AAE73688
C	45	30	3.2	1186	21	AAE98064

## ALIGNMENTS

RESULT	1
AD000121	
ID	AD000121 standard; DNA: 11933 BP.
XX	
AC	AD000121;
XX	
DT	31-JUL-2000 (first entry)
XX	
DE	Recombinant adeno associated vector construct, pAAV-F8-1.
XX	
XX	Recombinant Adeno Associated Vector; PAAV; PAAV-F8-1; human Factor VIII;
KM	hFVIII; hNF-3 albumin promoter; human elongation factor-1alpha; E1alpha;
KW	human growth hormone; hGH; Inverted terminal repeat; ITR; haemophilia;
KW	gene therapy; ds.
XX	
OS	Adeno associated virus.
XX	
PN	WO200023116-A1.
XX	
PD	27-APR-2000.
XX	
PE	19-OCT-1999; 99WO-US24495.
XX	
XX	20-OCT-1998; 98US-0104994.
PR	24-MAR-1999; 99US-0125974.
PR	30-JUL-1999; 99US-0364862.
XX	
PA	(AVIG-) AVIGEN INC.
XX	
PI	Couto LB, Colosi PC;
XX	
DR	WPI; 2000-339536/29.
XX	
PT	New recombinant adenovirus-associated vector, useful for gene therapy

2238 A; 2818 C; 2717 G; 3140 T; 0 other;

Matches	934	Conservative	0	Miscellaneous	22	21
Similarity	100.08					
Pred. NO.	4.9e-312					

DbpA  
11000 aaacgcgcaagcagatcttccaccggcttcacgaagtcatgtgta 60

[illegible]

11130 .....cgcgacgacgtctacgcatttgtagtcggttcttc 180

161 ttcgcgtatcgcctaattcagcacacaagaaagagcactgcttaaccagagctcagcnnac 340

241 tcttcacgattatcgactcaatgctcttaacctgttttcgaagatataaaas+cc

301 cgtatgcagcgtctaactattactctggaacatttcacacatcccaaac 11299

361 tctgcctgcgatqattgaagatccagagacata-  
-----cgcgcgaacacgllcccgggatgcattttgcagacctc 11359

421 qtaqtaagcccccctatattt  
11419

52235354555657585960616263646566676869707172737475767778798081828384858687888990919293949596979899100101102103104105106107108109110111112113114115116117118119120121122123124125126127128129130131132133134135136137138139140141142143144145146147148149150151152153154155156157158159160161162163164165166167168169170171172173174175176177178179180181182183184185186187188189190191192193194195196197198199200201202203204205206207208209210211212213214215216217218219220221222223224225226227228229230231232233234235236237238239240241242243244245246247248249250251252253254255256257258259260261262263264265266267268269270271272273274275276277278279280281282283284285286287288289290291292293294295296297298299300301302303304305306307308309310311312313314315316317318319320321322323324325326327328329330331332333334335336337338339340341342343344345346347348349350351352353354355356357358359360361362363364365366367368369370371372373374375376377378379380381382383384385386387388389390391392393394395396397398399400401402403404405406407408409410411412413414415416417418419420421422423424425426427428429430431432433434435436437438439440441442443444445446447448449450451452453454455456457458459460461462463464465466467468469470471472473474475476477478479480481482483484485486487488489490491492493494495496497498499500501502503504505506507508509510511512513514515516517518519520521522523524525526527528529530531532533534535536537538539540541542543544545546547548549550551552553554555556557558559560561562563564565566567568569570571572573574575576577578579580581582583584585586587588589590591592593594595596597598599600601602603604605606607608609610611612613614615616617618619620621622623624625626627628629630631632633634635636637638639640641642643644645646647648649650651652653654655656657658659660661662663664665666667668669670671672673674675676677678679680681682683684685686687688689690691692693694695696697698699700701702703704705706707708709710711712713714715716717718719720721722723724725726727728729730731732733734735736737738739740741742743744745746747748749750751752753754755756757758759760761762763764765766767768769770771772773774775776777778779780781782783784785786787788789790791792793794795796797798799800801802803804805806807808809810811812813814815816817818819820821822823824825826827828829830831832833834835836837838839840841842843844845846847848849850851852853854855856857858859860861862863864865866867868869870871872873874875876877878879880881882883884885886887888889890891892893894895896897898899900901902903904905906907908909910911912913914915916917918919920921922923924925926927928929930931932933934935936937938939940941942943944945946947948949950951952953954955956957958959960961962963964965966967968969970971972973974975976977978979980981982983984985986987988989990991992993994995996997998999100010011002100310041005100610071008100910101011101210131014101510161017101810191020102110221023102410251026102710281029103010311032103310341035103610371038103910401041104210431044104510461047104810491050105110521053105410551056105710581059106010611062106310641065106610671068106910701071107210731074107510761077107810791080108110821083108410851086108710881089109010911092109310941095109610971098109911001100110021003100410051006100710081009101010111012101310141015101610171018101910201021102210231024102510261027102810291030103110321033103410351036103710381039104010411042104310441045104610471048104910501051105210531054105510561057105810591060106110621063106410651066106710681069107010711072107310741075107610771078107910801081108210831084108510861087108810891090109110921093109410951096109710981099110011001100210031004100510061007100810091010101110121013101410151016101710181019102010211022102310241025102610271028102910301031103210331034103510361037103810391040104110421043104410451046104710481049105010511052105310541055105610571058105910601061106210631064106510661067106810691070107110721073107410751076107710781079108010811082108310841085108610871088108910901091109210931094109510961097109810991100110011002100310041005100610071008100910101011101210131014101510161017101810191020102110221

11480 gccaatttaccgaatttattgataagaatcatcatgcggtatataatgttgcacac 11520

11540 gtgcaatcatgctgctaacgtgtgacgcgcatlcaaataqtltgtctmattacacattt  
553acgacccccc 600

660

## RESULT 2

ID	standard; DNA; 11933 BP.
XX	

DT 04-SEP-2001 (first entry)  
XX

Human factor VII expressi

Recombinant adeno-associated virus (rAAV) partial sequence.

blood clotting disorder; gene therapy; haemophilia A; human; PAAV-F8-1; ds.

Chimeric - Adeno associated virus

Chimeric - Mus sp.  
Chimeric - Homo sapiens

Chimeric - *Oryctolagus cuniculus*.

WO200145510-A1

D  
X 28-JUN-2001.

21-DEC-2000; 2000WO-US34925

22-DEC-1999; 99US-0470618

(AVIG-) AVIGEN INC

Couto LB. Colocesi, 1901.

WPI: 2001-417055

Treatise 11

administering recombinant blood clot

coagulation factor

Claim 18; Fig 5; 90pp; English

The present invention relates to a method for treating a patient suffering from a blood disease.

administering a recombinant adeno-associated virus (rAAV) to a subject suffering from a blood clotting disorder. The method comprises

second recombinant adeno-associated virus containing the light chain of factor VIII and a sequence encoding the light chain of factor VIII (rAAV) comprising

useful in gene therapy for treating haemophilia. The rAAV vector is

particular humans. The rAAV vector provides high level and long term expression of biologically active proteins.

The present sequence is pAAV-P8-1 vector without the plasmid

CC backbone. It comprises HNF-3, mouse albumin promoter, a synthetic intron  
CC based on human elongation factor 1alpha (EF1alpha), and immunoglobulin G  
CC (IgG) intron sequences, B-domain deleted human factor VIII coding  
CC sequence, poly A signal based on rabbit beta globin sequence and AAV  
CC inverted terminal repeats (ITRs) at the ends. The vector encodes both the  
CC light and heavy chains of human factor VIII.

XX Sequence 1193 BP; 3258 A; 2818 C; 2717 G; 3140 T; 0 other;

Score 934; DB 22; Length 119337

Query Match	100.0%	Score	254	2	1
Best Local Similarity	100.0%	Pred. No.	4	9e-312	
Matches	934	Conservative	0	Mismatches	0
				Indels	0
				Gaps	0

[illegible][illegible]

RESULT	3
AAA11992	
ID AAA11992 standard; DNA; 37856 BP	

AC AAA11992;

AA  
DT 07-AUG-2000 (first entry)

XX encoding polyketide and hereropolyketide enzymes.

DE S. cerevisiae; heteropolyketide biosynthesis; mutasynthesis;  
XX immunosuppressant; antibiotic; antifungal;  
KW polyketide biosynthesis; heteropolyketide biosynthesis;  
mutasynthesis; immunosuppressant; antibiotic; antifungal;

plant-protection; ds.  
KW

XX  
OS Sorangium cellulosum.

XX FH	Key	Location/Qualifiers
		6100)
		3398

CD5	FT	FT
/*tag = a		
/*product =		
"ORF1-tRNA synthetase"		

```

/!note= "gtg start codon"
      /!start (6374-7111)

```

/\*tag= b "OPR2-monoxygenase"

complement (8433..9550)

ET  
ET  
ET

/product= "ORF3-aminotransferase"  
start="AGT start codon given in the specification"

FT	9855.11393
FT	9855.11393
CDS	9855.11393

```

ET      /product= "ORF4- tyrosin
FT      /start= "CTG start codon"

```

F1	12212..13658
F1T	+
CDS	+

```

E1 /product= "ORF5-3-OXoacyl acf 10000
FT /note= "ACC start codon"

```

FT	15374..19984
FT	/**tag= f
CDS	

```

      /product="ORFeo-polyketide
      30003  27889
      FT
      FT

```

```

      /*tag= 9 "ORF7-peptide s
      /product=

```

FT	28251..29400
CDS	/*tag= h

```

      /product= "UKF8-clauspep"
      complement (30040..31720)

```

```
/*tag= 1
/product= "ORF9-regulation"

```

```

/Note= "CGC stop coon
31982.32932

```

```

/*tag= ]
/product= "ORF10-transcr
FT

```

FT	33128..33613
CDS	/*tag= k

```

/ product= "OKHII-regula
/ note= "GTG start codon"

```

33661..340//  
/\*täq= 1

```

/produit= "ORF12-regulation
complement (35255..35616

```

```
/*tag= m
/product= "ORF13-transcr
```

```
complement (35/50...0023
/*tag= n
```

```

/product= "OKR14 crabs"
/note= "GTG start codon"

```

XX

PN DE19846493-A1.  
XX  
PD 13-APR-2000.  
XX  
PF 09-OCT-1998; 98DE-1046493.  
XX  
PR 09-OCT-1998; 98DE-1046493.  
XX  
PA (GBFB) GES BIOTECHNOLOGISCHE FORSCHUNG MBH;  
XX  
PI Beyer S, Mueller R;  
XX  
DR WPI: 2000-294101/26.  
XX  
PT DNA sequence coding for products involved in the biosynthesis of  
XX polyketide or heteropolysaccharide compounds, especially epothilone  
PS Claim 3; Page 20-33; 36pp; German.  
XX  
CC This invention describes a novel DNA sequence (I) whose expression  
CC products effect or are involved in the enzymatic biosynthesis,  
CC mutasynthesis or partial synthesis of polyketide or heteropolysaccharide  
CC compounds (II). (I) can be inserted into an expression vector and used  
CC to transform or transfect prokaryotic or eukaryotic cells with the aim  
CC of obtaining strains that produce large amounts of polyketide or  
CC heteropolysaccharide compounds, especially epothilones, which have cytotoxic  
CC and/or immunosuppressant and antibiotic and antifungal activities and  
CC are useful as plant-protection agents. This sequence represents the DNA  
CC sequence isolated from *Sorangium cellulosum* which is described in the  
XX method of the invention.  
XX  
SQ Sequence 37856 BP; 5655 A; 13666 C; 12913 G; 5622 T; 0 other;  
  
Query Match  
Best Local Similarity 3.9%; Score 36.6; DB 21; Length 37856;  
Matches 63; Conservative 0; Mismatches 44; Indels 0; Gaps 0;  
  
Qy 756 tctcgacaactcccccagagctgttggcaatatacaccgacgctgttaccat 815  
Db 25067 ttcgacacaactctgtcctcaagaagtgctacgctctacgagcgacgacgacacgt 25126  
Qy 816 ggaactctctgacatcttgcctcccgagctcgagcagcagcgaataa 862  
Db 25127 gggacatctcgctgagctcgagcagcagcggcctcagcgagatta 25173  
  
RESULT 4  
AAC37411  
ID AAC37411 standard; DNA; 1063 BP.  
XX  
AC AAC37411;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 17282.  
XX  
KW Hybridisation assay; genetic mapping; gene expression control;  
KW Protein identification; signal transduction pathway;  
KW metabolic pathway; promoter; termination sequence; ss.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EPI033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132485.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 18-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137528.  
PR 07-JUN-1999; 99US-0137502.  
PR 08-JUN-1999; 99US-0137724.  
PR 10-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138504.  
PR 14-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139452.  
PR 18-JUN-1999; 99US-0139452.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139465.  
PR 18-JUN-1999; 99US-0139750.  
PR 21-JUN-1999; 99US-0139763.  
PR 22-JUN-1999; 99US-0139817.  
PR 23-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 24-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.

PR	16-JUL-1999	9905-0144306
PR	16-JUL-1999	9905-0144333
PR	19-JUL-1999	9905-0144331
PR	19-JUL-1999	9905-0144332
PR	19-JUL-1999	9905-0144333
PR	19-JUL-1999	9905-0144334
PR	19-JUL-1999	9905-0144335
PR	19-JUL-1999	9905-0144332
PR	20-JUL-1999	9905-0144632
PR	20-JUL-1999	9905-0144684
PR	20-JUL-1999	9905-0144814
PR	21-JUL-1999	9905-0145086
PR	21-JUL-1999	9905-0145088
PR	21-JUL-1999	9905-0145083
PR	22-JUL-1999	9905-0145087
PR	22-JUL-1999	9905-0145089
PR	22-JUL-1999	9905-0145192
PR	23-JUL-1999	9905-0145145
PR	23-JUL-1999	9905-0145148
PR	23-JUL-1999	9905-0145274
PR	26-JUL-1999	9905-0145276
PR	27-JUL-1999	9905-0145213
PR	27-JUL-1999	9905-0145518
PR	27-JUL-1999	9905-0145919
PR	28-JUL-1999	9905-0145951
PR	28-JUL-1999	9905-0146386
PR	02-AUG-1999	9905-0146388
PR	02-AUG-1999	9905-0146389
PR	02-AUG-1999	9905-0147038
PR	03-AUG-1999	9905-0147204
PR	04-AUG-1999	9905-0147302
PR	04-AUG-1999	9905-0147392
PR	05-AUG-1999	9905-0147260
PR	05-AUG-1999	9905-0147303
PR	06-AUG-1999	9905-0147416
PR	06-AUG-1999	9905-0147435
PR	09-AUG-1999	9905-0147933
PR	09-AUG-1999	9905-0148177
PR	10-AUG-1999	9905-0148319
PR	11-AUG-1999	9905-0148341
PR	12-AUG-1999	9905-0148655
PR	13-AUG-1999	9905-0148684
PR	13-AUG-1999	9905-0149368
PR	16-AUG-1999	9905-0149375
PR	17-AUG-1999	9905-0149426
PR	18-AUG-1999	9905-0149622
PR	20-AUG-1999	9905-0149723
PR	20-AUG-1999	9905-0149929
PR	20-AUG-1999	9905-0149930
PR	22-AUG-1999	9905-0149930
PR	23-AUG-1999	9905-0150566
PR	25-AUG-1999	9905-0150884
PR	26-AUG-1999	9905-0151065
PR	27-AUG-1999	9905-0151066
PR	27-AUG-1999	9905-0151080
PR	30-AUG-1999	9905-0151300
PR	30-AUG-1999	9905-0151438
PR	01-SEP-1999	9905-0151930
PR	01-SEP-1999	9905-0152363
PR	07-SEP-1999	9905-0153070
PR	10-SEP-1999	9905-0153758
PR	13-SEP-1999	9905-0154018
PR	15-SEP-1999	9905-0154039
PR	16-SEP-1999	9905-0154779
PR	22-SEP-1999	9905-0155139
PR	23-SEP-1999	9905-0155386
PR	24-SEP-1999	9905-0155659
PR	28-SEP-1999	9905-0156438
PR	28-SEP-1999	9905-0156586
PR	04-OCT-1999	9905-0157117
PR	05-OCT-1999	9905-0157753
PR	07-OCT-1999	9905-0157855
PR	07-OCT-1999	9905-0158029

	Query Match	3.5%	Score 33;	DB 21;	Length 1063;
	Best Local Similarity	52.6%;	Pred. No. 0.86;	Indels	0; Gaps
	Matches 72;	Conservative	0;	Mismatches	65;
Oy	466 ttaaccatcggttgagtcacaatttaccgaattttatccaataatgaatcaatatcgcggtt	525			
Dd	866 ttcccaactcatctgaaataatgatctaacttatattgaacaaacagcacatccctctct	925			
Oy	526 aaatagtgtccatccgcgtggagaatcagtcgtcctaagctggcaggcgcataaatgttgtc	585			
Dd	926 tgtaaagatgcataactgaataagaacagatgacaacatttgatctctctcttgttgtt	985			
Oy	586 tgcgattgacctctctt	602			
Dd	986 tgccttatgctctcat	1002			

RESULT      5

ID	AAF28541	standard;	DNA;	49617	BP.
XX	AAF28541:				
AC					
XX					
DT	04-APR-2001	(first entry)			
XX					
DE	Genomic fragment #28.				
XX					
KW	Genomic library; bacteria; human upper airway; otitis media; sinusitis;				
KM	Bronchopulmonary; endocarditis; meningitis; ss.				
OS	Moraxella catarrhalis.				
XX					
PN	WO200078968-A2.				
XX					
PD	28-DEC-2000.				
XX					
PF	16-JUN-2000; 2000MO-US16649.				
XX					
PR	18-JUN-1999; 99US-0140121.				
XX					
PA	(INCY-) INCYTE GENOMICS INC.				
XX					

PI Lagace RE, Patterson C; Berg KL;  
 DR WPI: 2001-041427/05.  
 XX  
 PT Genomic library for identifying diagnostic and therapeutic  
 PT compositions, and for identifying virulence factors, regulatory  
 PT elements and drug targets, comprises Moraxella catarrhalis nucleic  
 acids -  
 XX  
 PS Claim 1: Page 235-247; 545pp; English.

CC The present invention relates to a Moraxella catarrhalis genomic library  
 CC comprising of a combination of 41 nucleic acid molecules (see  
 CC AAF28514-AAF28554). The library has a number of uses described in the  
 CC specification e.g. is useful for identifying diagnostic and therapeutic  
 CC compositions. M. catarrhalis (Branhamella catarrhalis) is a large  
 CC aerobic gram-negative diplococcus, normally found among the bacterial  
 CC flora of human upper airways. M. catarrhalis is known to cause acute,  
 CC localised infections such as otitis media, sinusitis and bronchopulmonary  
 CC infection and life-threatening, systemic diseases including endocarditis  
 CC and meningitis.  
 XX  
 SQ Sequence 49617 BP; 14572 A; 9853 C; 11026 G; 14166 T; 0 other;

Query Match  
 Best Local Similarity 3.4%; Score 32; DB 22; Length 49617;  
 Matches 68; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

OY 517 catgcgtatattgttcgctcgtggaatcattgctgtaacgtgaccgcatcaca 576  
 DB 11552 caccgcttgatgattgctcacaagcttgataatctgctcgtgaagaacacacacgctccat 11621  
 OY 577 aattgtctgcatgactctctctgtgcatgcacccagagcgctacatacagcg 656  
 DB 11622 tgggtcgtacagagtgactcttaggcgctactctgcaccacagcgctgatacaca 11681  
 OY 637 cttaacag 644  
 DB 11682 cggatgag 11689

RESULT 6  
 AA112818/c  
 ID AA112818 standard; DNA; 469 BP.  
 XX  
 AC AA112818;  
 XX  
 DT 12-OCT-2001 (first entry)  
 DE Probe #2751 for gene expression analysis in human cervical cell sample.  
 KW Probe: human; microarray; gene expression; cervical epithelial cell;  
 KW cervical cancer; ss.  
 XX  
 OS Homo sapiens.  
 PN W0200157278-A2.  
 PD 09-AUG-2001.  
 PF 30-JAN-2001; 2001WO-US00670.  
 XX  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 DR WPI: 2001-488901/53.  
 XX  
 PT Human genome-derived single exon nucleic acid probes useful for  
 PT analyzing gene expression in human cervical epithelial cells -  
 XX  
 PS Claim 25; SEQ ID No 2751; 487pp; English.

CC The present invention relates to human single exon nucleic acid probes  
 CC (SENPs). The present sequence is one such probe. The SENPs are derived  
 CC from human HeLa cells. The SENPs can be used to produce a single exon  
 CC microarray, which can be used for measuring human gene expression in a  
 CC sample derived from human cervical epithelial cells. By measuring gene  
 CC expression, the probes are therefore useful in grading and/or staging  
 CC of diseases of the cervix, notably cervical cancer.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pcl\_sequences.  
 XX  
 SQ Sequence 469 BP; 125 A; 159 C; 88 G; 97 T; 0 other;

Query Match  
 Best Local Similarity 3.4%; Score 31.8; DB 22; Length 469;  
 Matches 54; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

OY 98 cagcagtgctcagctcagctgattggtgtgctgtctgctgagcgatgacgctgta 157  
 DB 115 CAGCAGACACACAGAGATTGAGTGGGGGCTGTGCTGTGCTGTGAGGACAGTGGCAATGC 56  
 OY 158 cgcattgtgacccggtctctctcgcta 188  
 DB 55 AAACCTGGGCTTCTGCGAGAGGCTTCTCGCA 25

RESULT 7  
 AA134173/c  
 ID AA134173 standard; DNA; 469 BP.  
 XX  
 AC AA134173;  
 XX  
 DT 17-OCT-2001 (first entry)  
 DE Probe #2859 used to measure gene expression in human placenta sample.  
 KW Probe: microarray; human; placenta; antenatal diagnosis;  
 KW genetic disorder; ss.  
 XX  
 OS Homo sapiens.  
 PN W0200157272-A2.  
 PD 09-AUG-2001.  
 PF 30-JAN-2001; 2001WO-US00663.  
 XX  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 DR WPI: 2001-488901/53.  
 XX  
 PT Human genome-derived single exon nucleic acid probes useful for  
 PT analyzing gene expression in human placenta -  
 XX



```

XX Claim 25; SEQ ID No 2859; 654pp: English.
XX The present invention relates to single exon nucleic acid probes (SENPs).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders.
XX Sequence 469 BP; 125 A; 159 C; 88 G; 97 T; 0 other;
SQ
Query Match          3.4%; Score 31.8; DB 22; Length 469;
Best Local Similarity 59.3%; Pred. No. 1.5;
Matches   54; Conservative 0; Mismatches    37; Indels    0; Gaps    0;
Oy      98 cagaacagtgcgtaagcggcttcgttggttgcctgctgcgcgcggaatcacctgta 157
         ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      115 CAGACAGAAGCGCAGAGATTGAGTGGGGGTCTTGCTTCGACAGCAAGTGCACAATC 56
         ||||| | | | | | | | | | | | | | | | | | | | | | | | |
Oy      158 ccgatctgtgtatccgattccgattcgcctcgcta 188
         ||||| | | | | | | | | | | | | | | | | | | | | | | | |
Db      55 AAACCTGGGTCTTCGCGAAGAGCTTCCTGGA 25
         ||||| | | | | | | | | | | | | | | | | | | | | | | | |

RESULT      8
AAI02735/c
ID   AAI02735 standard; DNA; 469 BP.
AC     AAI02735.
DT     09-OCT-2001 (first entry) .
DE     Probe #2726 used to measure gene expression in human breast sample.
KW     Probe; human; breast disease; breast cancer; development disorder; ss;
KM     inflammatory disease; proliferative breast disease; non-carcinoma tumour
OS     Homo sapiens.
PN     WO200157270-A2.
PD     09-AUG-2001.
PE     29-JAN-2001; 2001WO-US00661.
PF     04-FEB-2000; 2000US-0180312.
PR     26-MAY-2000; 2000US-0207456.
PR     30-JUN-2000; 2000US-0608408.
PR     03-AUG-2000; 2000US-0632366.
PR     21-SEP-2000; 2000US-0234687.
PR     27-SEP-2000; 2000US-0236359.
PR     04-OCT-2000; 2000GB-0024263.
PA     (MOLE-) MOLECULAR DYNAMICS INC.
PI     Penn SG, Hanzel DK, Chen W, Rank DR;
PX     WPI; 2001-476286/51.
PT     Novel single exon nucleic acid probe used to measuring gene expression
       in a human breast -
PS     Claim 25; SEQ ID No 2726; 322pp: English.
XX The present invention relates to novel single exon nucleic acid probes.
CC The present sequence is one such probe. The probes are useful for
CC measuring human gene expression in a human breast sample, where the probe
CC hybridises at high stringency to a nucleic acid expressed in the human
CC breast. The probes are useful for predicting, diagnosing, grading,
CC staging, monitoring and prognosing diseases of the human breast,
CC particularly those diseases with polygenic aetiology. The diseases
CC include: breast cancer, disorders of development, inflammatory diseases

```

CC	of the breast, fibrocystic changes, proliferative breast disease and
CC	non-carcinoma tumours.
CC	Note: The sequence data for this patent did not form part of the printed
CC	specification, but was obtained in electronic format directly from WIPO
CC	at ftp.wipo.int/pub/published_pct_sequences.
XX	
SO	Sequence 469 BP; 125 A; 159 C; 88 G; 97 T; 0 other;
Qy	Query Match 3.4%; Score 31.8; DB 22; Length 469; Best Local Similarity 59.3%; Pred.No.1.5; Matches 54; Conservative 0; Mismatches 37; Indels 0; Gaps 0;
Dy	98 cagcaggctgctacgcgttcagtttggtggctgctggcgccgagtgaacctgta 157  Dy 115 CAGACGACGACGAGATTGATGGCGGTCTGTGTTGTCGACGACAAGTGCAGAATC 56 
Oy	158 cgcatctggtgatcccgattctctcttcgtagta 188   Db 55 AACCTTGGTGTCTGTGGGAAGGCTCTCGA 25
RESULT 9	
AAl18836/C	
ID	AAl18836 standard; DNA: 542 BP.
XX	
AC	AAl18836;
DT	12-OCT-2001 (first entry)
XX	
DE	Probe #8769 for gene expression analysis in human cervical cell sample. Probe: human: microarray; gene expression: cervical epithelial cell};
XX	
KM	cervical cancer; ss.
XX	
OS	Homo sapiens.
XX	
PN	MO200157278-A2.
PD	
XX	
PF	30-JAN-2001; 2001WO-US00670.
XX	
PR	04-FEB-2000; 2000US-0180312.
PR	26-MAY-2000; 2000US-0207456.
PR	30-JUN-2000; 2000US-0608408.
PR	03-AUG-2000; 2000US-0632366.
PR	21-SEP-2000; 2000US-0234687.
PR	27-SEP-2000; 2000US-0236359.
PR	04-OCT-2000; 2000GB-0024263.
XX	
PA	(MOLE-) MOLECULAR DYNAMICS INC.
PI	
PT	Penn SG, Hanzel DK, Chen W, Rank DR;
XX	
DR	WPI: 2001-488901/53.
XX	
XX	Human genome-derived single exon nucleic acid probes useful for analyzing gene expression in human cervical epithelial cells -
XX	Claim 25; SEQ ID No 8769; 487bp; English.
XX	The present invention relates to human single exon nucleic acid probes (SENPs). The present sequence is one such probe. The SENPs are derived from human HeLa cells. The SENPs can be used to produce a single exon microarray, which can be used for measuring human gene expression in a sample derived from human cervical epithelial cells. By measuring gene expression, the probes are therefore useful in grading and/or staging of diseases of the cervix, notably cervical cancer. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX	



Db 374 WWWWWW 315  
Qy 549 cagcgcgtacagtggtaccgcatcaaatgtgtcgcgattgactctctt 603  
Db 314 WWWWWW 260

## RESULT 12

AAF58254/c  
ID AAF58254 standard; DNA; 936 BP.

AC AAF58254;

DT 24-APR-2001 (first entry)

DE Oligonucleotide D1875.

XX Electron-transfer group; ETM; mismatch; genotyping;

KW gene expression; ss.

OS Synthetic.

PN WO200107665-A2.

PD 01-FEB-2001.

PF 26-JUL-2000; 2000WO-US20476.

PR 26-JUL-1999; 99US-0145655.

PR 17-MAR-2000; 2000US-0190259.

PA (CLIN-) CLINICAL MICRO SENSORS INC.

PI Umek RM.

DR WPI: 2001-159728/16.

XX Nucleic acids containing electron-transfer group, useful as labels in

PT hybridization assays, e.g. for genotyping, allowing repeat analyses on

PT a single surface

XX Example 6; Page 127; 159pp; English.

PS The present invention relates to a composition comprising two nucleic

XX acids each containing an electron-transfer group (ETM) having

CC different redox potentials. The invention is used for electronic

CC detection of nucleic acids, especially of substitutions (mismatches)

CC and single-nucleotide polymorphisms, e.g. for genotyping,

CC monitoring gene expression.

CC Sequence 936 BP; 4 A; 144 C; 7 G; 5 T; 776 other;

XX

XX

XX

XX

XX

Qy 489 taccgaatttcaataagatcaatgcgttgatgacatcggtgcaat 548  
Db 374 WWWWWW 315  
Qy 549 cagcgcgtacagtggtaccgcatcaaatgtgtcgcgattgactctctt 603  
Db 314 WWWWWW 260

## RESULT 13

AAF58257/c  
ID AAF58257 standard; DNA; 936 BP.

AC AAF58257;

DT 24-APR-2001 (first entry)

DE Oligonucleotide D1954.

XX Electron-transfer group; ETM; mismatch; genotyping;

KW gene expression; ss.

OS Synthetic.

PN WO200107665-A2.

PD 01-FEB-2001.

PF 26-JUL-2000; 2000WO-US20476.

PR 26-JUL-1999; 99US-0145655.

PR 17-MAR-2000; 2000US-0190259.

PA (CLIN-) CLINICAL MICRO SENSORS INC.

PI Umek RM.

DR WPI: 2001-159728/16.

XX Nucleic acids containing electron-transfer group, useful as labels in

PT hybridization assays, e.g. for genotyping, allowing repeat analyses on

PT a single surface

XX Example 6; Page 127; 159pp; English.

PS The present invention relates to a composition comprising two nucleic

XX acids each containing an electron-transfer group (ETM) having

CC different redox potentials. The invention is used for electronic

CC detection of nucleic acids, especially of substitutions (mismatches)

CC and single-nucleotide polymorphisms, e.g. for genotyping,

CC monitoring gene expression.

CC Sequence 936 BP; 5 A; 142 C; 7 G; 6 T; 776 other;

XX

XX

XX

XX

XX

Db 434 WWWWWWTTAAGCWWWWW 375  
Qy 489 taccattatcaatgaatcatcgcgttaattgttgcacccgtgcaat 548  
Db 374 WWWWWWTTAAGCWWWWW 375  
Qy 549 catgtgttaacgtgtgacgcacatcaaatgttgcgtgactctctt 603  
Db 314 WWWWWWTTAAGCWWWWW 260

RESULT 14  
ID AAF58259 standard; DNA: 936 BP.  
AC AAF58259:  
DF 24-APR-2001 (first entry)  
DE Oligonucleotide D2004.  
KM Electron-transfer group; ETM; mismatch: genotyping;  
KW gene expression; ss.  
XX  
OS Synthetic.  
PN WO200107665-A2.  
PD 01-FEB-2001.  
PE 26-JUL-2000; 2000WO-US20476.  
PR 26-JUL-1999; 99US-0145695.  
PR 17-MAR-2000; 2000US-0190259.  
PA (CLIN-) CLINICAL MICRO SENSORS INC.  
XX  
PI Umek RM;  
DR WPI: 2001-159728/16.  
XX

XX Nucleic acids containing electron-transfer group, useful as labels in  
PT hybridization assays, e.g. for genotyping, allowing repeat analyses on  
a single surface  
XX

PS Example 6; Page 128; 159pp; English.

CC The present invention relates to a composition comprising two nucleic  
CC acids each containing an electron-transfer group (ETM) having  
CC different redox potentials. The invention is used for electronic  
CC detection of nucleic acids, especially of substitutions (mismatches)  
CC and single-nucleotide polymorphisms, e.g. for genotyping,  
CC monitoring gene expression.  
XX

SO Sequence 936 BP; 6 A; 138 G; 8 G; 8 T; 776 other;

Query Match 3.4%; Score 31.8; DB 22; Length 936;  
Best Local Similarity 1.7%; Pred. No. 2.1;  
Matches 6; Conservative 196; Mismatches 153; Indels 0; Gaps 0;

Qy 249 atatacgaacatgcttactgttgcgtatataaaatccgaacgttatgc 308  
Db 614 WWWWWWTTAAGCWWWWW 555  
Qy 309 agccttaactatcctgcgaactgttgcgtatgcatttgcagacctctgcctg 368  
Db 554 WWWWWWTTAAGCWWWWW 495  
Qy 369 cgaatgttgagttccagcgcatacgtgcaagtgacccaactaggcggaatcgtagtaag 428  
Db 494 GWWWWWTTAAGCWWWWW 435

Qy 429 cgcgcctcttctactcactacacagcgaatcaaccatcgttgatcaat 488  
Db 434 WWWWWWTTAAGCWWWWW 375  
Qy 489 taccattatcaatgaatcatcgcgttaattgttgcacccgtgcaat 548  
Db 374 WWWWWWTTAAGCWWWWW 375  
Qy 549 catgtgttaacgtgtgacgcacatcaaatgttgcgtgactctctt 603  
Db 314 WWWWWWTTAAGCWWWWW 260

RESULT 15  
ID AAF58262 standard; DNA: 936 BP.  
AC AAF58262:  
DF 24-APR-2001 (first entry)  
DE Oligonucleotide D2007.  
KM Electron-transfer group; ETM; mismatch: genotyping;  
KW gene expression; ss.  
XX  
OS Synthetic.  
PN WO200107665-A2.  
PD 01-FEB-2001.  
PE 26-JUL-2000; 2000WO-US20476.  
PR 26-JUL-1999; 99US-0145695.  
PR 17-MAR-2000; 2000US-0190259.  
PA (CLIN-) CLINICAL MICRO SENSORS INC.  
XX  
PI Umek RM;  
DR WPI: 2001-159728/16.  
XX

XX Nucleic acids containing electron-transfer group, useful as labels in  
PT hybridization assays, e.g. for genotyping, allowing repeat analyses on  
a single surface  
XX

PS Example 6; Page 128; 159pp; English.

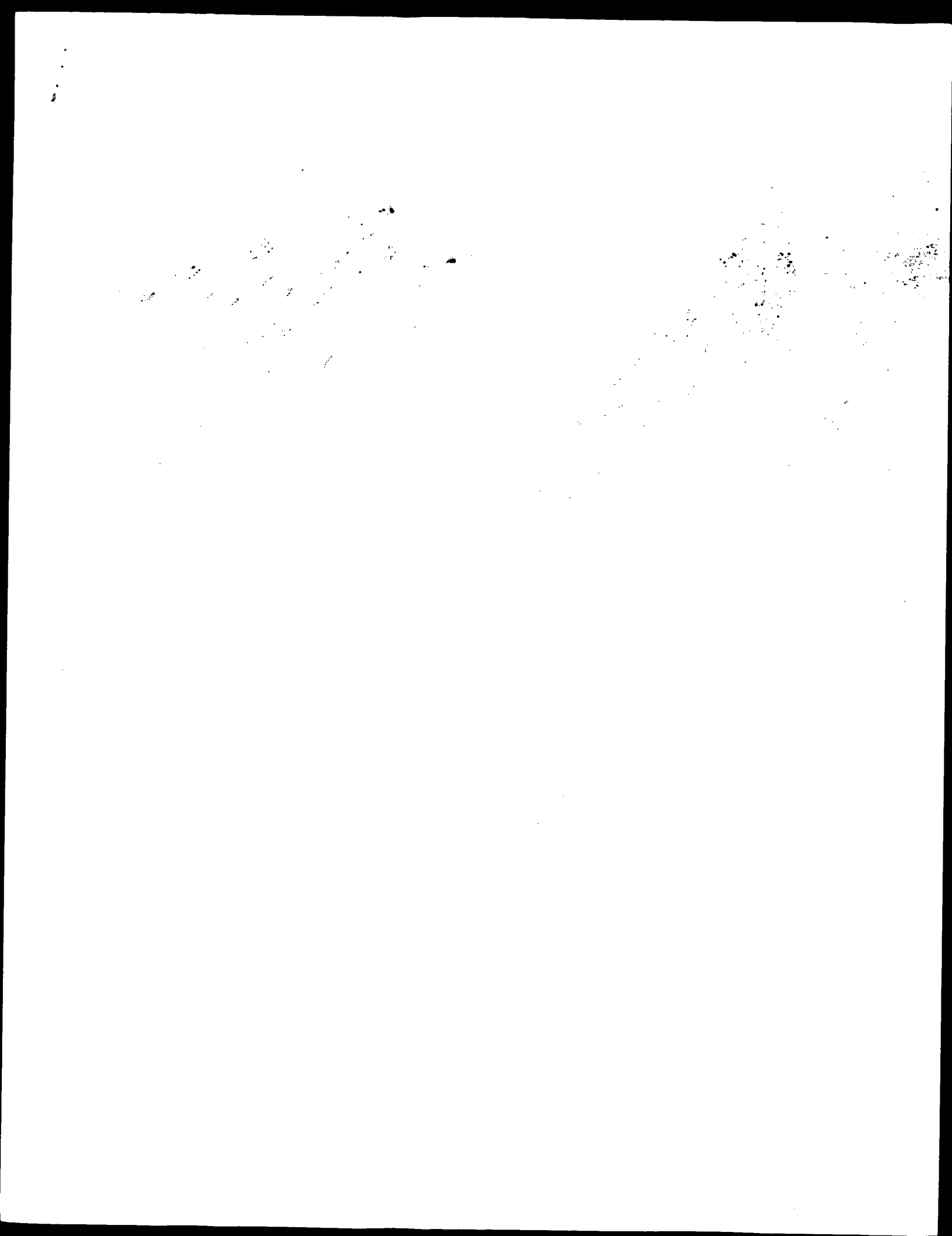
CC The present invention relates to a composition comprising two nucleic  
CC acids each containing an electron-transfer group (ETM) having  
CC different redox potentials. The invention is used for electronic  
CC detection of nucleic acids, especially of substitutions (mismatches)  
CC and single-nucleotide polymorphisms, e.g. for genotyping,  
CC monitoring gene expression.  
XX

SO Sequence 936 BP; 5 A; 139 G; 10 G; 6 T; 776 other;

Query Match 3.4%; Score 31.8; DB 22; Length 936;  
Best Local Similarity 1.7%; Pred. No. 2.1;  
Matches 6; Conservative 196; Mismatches 153; Indels 0; Gaps 0;

Qy 249 atatacgaacatgcttactgttgcgtatataaaatccgaacgttatgc 308  
Db 614 WWWWWWTTAAGCWWWWW 555  
Qy 309 agccttaactatcctgcgaactgttgcgtatgcatttgcagacctctgcctg 368  
Db 554 WWWWWWTTAAGCWWWWW 495  
Qy 369 cgaatgttgagttccagcgcatacgtgcaagtgacccaactaggcggaatcgtagtaag 428  
Db 494 GWWWWWTTAAGCWWWWW 435







Db 11240 tcttcacgattatcagctcaatgctcttaccgcttgcgagatataaaatccgaac 11299  
Qy 301 cgtatgcaagccttaactaactggaactgttctcgatgtgacttttgaagctc 360  
Db 11300 cgtatgcaagccttaactaactggaactgttctcgatgtgacttttgaagctc 360  
Qy 361 tctgctgcaatggttggatgctcgaagatgctgaagtgcaactgcttgcgaactc 11359  
Db 11360 tctgctgcaatggttggatgctcgaagatgctgaagtgcaactgcttgcgaactc 11359  
Qy 421 gtagtaagcgcgcctcttctcactcactcacaagagcgaaatgaagcgaatcg 11419  
Db 11420 gtagtaagcgcgcctcttctcactcactcacaagagcgaaatgaagcgaatcg 11419  
Qy 481 gtaattaccacatttattatcaatgaatgcaatcagcgttaaatatgttgcacac 11479  
Db 11480 gtaattaccacatttattatcaatgaatgcaatcagcgttaaatatgttgcacac 11479  
Qy 541 gtagcaatcagctgcttaacgctgtagcgcctcaaaatgttctcgaatgctcctc 11539  
Db 11540 gtagcaatcagctgcttaacgctgtagcgcctcaaaatgttctcgaatgctcctc 11539  
Qy 601 ttgtgcaatgcaacacacagagcgctacagcgcttaacagctgtagcgaagtgag 11599  
Db 11600 ttgtgcaatgcaacacacagagcgctacagcgcttaacagctgtagcgaagtgag 11599  
Qy 661 ttgtgcaatgcttggatgctacgctgtagcgaagtgagctgtagcgaatcct 720  
Db 11660 ttgtgcaatgcttggatgctacgctgtagcgaagtgagctgtagcgaatcct 720  
Qy 721 tgaatagcgaagccttgccttctcgaactcttctcgaacacacacacacacac 11719  
Db 11720 tgaatagcgaagccttgccttctcgaactcttctcgaacacacacacacacac 11719  
Qy 781 tttttgcaatcaac 11779  
Db 11780 tttttgcaatcaac 11779  
Qy 841 cgtcgcaatcagcgaatgaatcgcgaatgagcgaatgtagcgaatgtagcgaatc 840  
Db 11840 cgtcgcaatcagcgaatgaatcgcgaatgagcgaatgtagcgaatgtagcgaatc 840  
Qy 901 tgccttagatttctcctcaagctgcccctgcag 934  
Db 11900 tgccttagatttctcctcaagctgcccctgcag 934

## RESULT 2

Sequence 13, Application US/09364862  
Patent No. 6221349  
GENERAL INFORMATION:  
APPLICANT: Couto, Linda B.  
TITLE OF INVENTION: ADENO-ASSOCIATED VECTORS FOR EXPRESSION OF FACTOR VIII  
TITLE OF INVENTION: BY TAKEN  
FILE REFERENCE: AVIGEN-03743  
CURRENT APPLICATION NUMBER: US/09/364, 862  
EARLIER FILING DATE: 1999-07-30  
EARLIER APPLICATION NUMBER: 60/125, 974  
EARLIER FILING DATE: 1999-03-24  
EARLIER APPLICATION NUMBER: 60/104, 994  
NUMBER OF SEQ ID NOS: 14  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO: 13  
LENGTH: 11933  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic

US-09-364-862-13

Query Match 100.0%; Score 934; DB 4; Length 11933;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 934; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 aaagcgagaggttctgtagcgaactcctgcccgccttcaagagtgatgta 60  
Db 11000 aaagcgagaggttctgtagcgaactcctgcccgccttcaagagtgatgta 60  
Qy 61 aaagcgagaggttctgtagcgaactcctgcccgccttcaagagtgatgta 11059  
Db 11060 aaagcgagaggttctgtagcgaactcctgcccgccttcaagagtgatgta 11059  
Qy 121 tgggttctgctgttctgtagcgaactcctgcccgccttcaagagtgatgta 11119  
Db 11120 tgggttctgctgttctgtagcgaactcctgcccgccttcaagagtgatgta 11119  
Qy 181 ttcggtatcgttcaatcagcaacagagagcactgtagcgaactcctgcccgc 11179  
Db 11180 ttcggtatcgttcaatcagcaacagagagcactgtagcgaactcctgcccgc 11179  
Qy 241 tctcaagattacgactcaatgcttcttactctgttgcagatataaaatccgaac 11239  
Db 11240 tctcaagattacgactcaatgcttcttactctgttgcagatataaaatccgaac 11239  
Qy 301 cgtatgcaagccttaactaactggaactgttctcgatgtgacttttgaagctc 11299  
Db 11300 cgtatgcaagccttaactaactggaactgttctcgatgtgacttttgaagctc 11299  
Qy 361 tctgctgcaatggttggatgctcgaagatgctgaagtgcaactgtagcgaatc 11359  
Db 11360 tctgctgcaatggttggatgctcgaagatgctgaagtgcaactgtagcgaatc 11359  
Qy 421 gtagtaagcgcgcctcttctcactcactcacaagagcgaaatgaagcgaatc 11419  
Db 11420 gtagtaagcgcgcctcttctcactcactcacaagagcgaaatgaagcgaatc 11419  
Qy 481 gtaattaccacatttattatcaatgaatgcaatcagcgttaaatatgttgcacac 11479  
Db 11480 gtaattaccacatttattatcaatgaatgcaatcagcgttaaatatgttgcacac 11479  
Qy 541 gtagcaatcagctgcttaacgctgtagcgcctcaaaatgttctcgaatgctcctc 11539  
Db 11540 gtagcaatcagctgcttaacgctgtagcgcctcaaaatgttctcgaatgctcctc 11539  
Qy 601 ttgtgcaatgcaacacacagagcgctacagcgcttaacagctgtagcgaagtgag 11599  
Db 11600 ttgtgcaatgcaacacacagagcgctacagcgcttaacagctgtagcgaagtgag 11599  
Qy 661 ttgtgcaatgcttggatgctacgctgtagcgaagtgagctgtagcgaatcct 720  
Db 11660 ttgtgcaatgcttggatgctacgctgtagcgaagtgagctgtagcgaatcct 720  
Qy 721 tgaatagcgaagccttgccttctcgaactcttctcgaacacacacacacacac 11719  
Db 11720 tgaatagcgaagccttgccttctcgaactcttctcgaacacacacacacacac 11719  
Qy 781 tttttgcaatcaac 11779  
Db 11780 tttttgcaatcaac 11779  
Qy 841 cgtcgcaatcagcgaatgaatcgcgaatgagcgaatgtagcgaatgtagcgaatc 840  
Db 11840 cgtcgcaatcagcgaatgaatcgcgaatgagcgaatgtagcgaatgtagcgaatc 840  
Qy 901 tgccttagatttctcctcaagctgcccctgcag 934  
Db 11900 tgccttagatttctcctcaagctgcccctgcag 934

## RESULT 3



Med Jan 9 14:38:17 2002

US-09-740-211-13\_COPY\_11000\_11933.ini

US-08-232-463-14 Application US/08232463

Sequence 14, Application US/08232463  
Patent No. 560367  
GENERAL INFORMATION:  
APPLICANT: DORNER, F.  
INVENTOR: SCHEITLINGER, F.  
ATTORNEY: FALKNER, F. G.  
TITLE OF INVENTION: 52  
TITLE OF SEQUENCE: 52  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE: Foley & Lardner  
ADDRESS: 1800  
STREET: Alexandria  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22313  
COMPUTER READABLE FORM: disk  
MEDIUM TYPE: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0; Version #1.25  
CURRENT APPLICATION NUMBER: US/08/232,463  
APPLICATION DATE: 435  
CLASSIFICATION DATA: US/07/935,313  
FILING DATE: 435  
PRIOR APPLICATION NUMBER: EP 91 114 300.6  
FILING DATE: 26-AUG-1991  
APPLICATION DATE: 26-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: STEPHEN A. BENT, STEPHEN A. BENT  
RESIDENCE/DOCKING INFORMATION:  
REFERENCE/DOCKING INFORMATION:  
TELEPHONE: (703) 683-4109  
TELEFAX: (703) 683-4109  
TELEX: 899118  
FILING DATE: 899118  
INFORMATION FOR CITERISTICS:  
INFORMATION FOR CITERISTICS:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7218 base pairs  
TYPE: nucleic acid  
STRANDNESS: linear  
TOPOLOGY: SOURCE:  
IMMEDIATE PUSAPT-FIS  
CLONE: PT34PT-FIS

US-08-232-463-14  
4.0%: Score 37.6; DB 1; Length 7218; Gaps 0;  
Best Local 7; Conservative 2.7%; 152; Mismatches 101; Indels 0;  
Matches  
Query Match similarity 2.7%; 152; Mismatches 101; Indels 0;  
Best Local 7; Conservative 2.7%; 152; Mismatches 101; Indels 0;  
Matches  
DB 670 gttggagatagc...  
DB 1058 gcttcgcatg...  
DB 730 gacgccttgac...  
DB 1118 yyyyyyy...  
DB 790 aatatacgc...  
DB 1178 yyyyyyy...  
DB 850 actacgc...  
DB 1238 yyyyyyy...  
DB 910 attcctcaagc...  
DB 1298 yyyyyyy...

RESULT 4 Application US/09045301A

US-09-045-301-1  
Sequence 1, Application US/09045301A  
Patent No. 6265388  
GENERAL INFORMATION:  
APPLICANT: Felt, Karen A.  
INVENTOR: Olson, Karen A.  
ATTORNEY: Olson, Karen A.  
TITLE OF INVENTION: 10498/05286  
TITLE OF SEQUENCE: 10498/05286  
FILE REFERENCE: 10498/05286  
CURRENT APPLICATION NUMBER: 60/041182  
CURRENT FILING DATE: 1997-03-21  
EARLIER FILING DATE: 1997-03-21  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: Patent Ver. 2.0  
SEQ ID NO: 1  
SEQ LENGTH: 4668  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURES: CDS  
NAME/KEY: (1809)..(2252)  
LOCATION: (1809)..(2252)  
US-09-045-301-1  
3.4%: Score 31.6; DB 4; Length 4668; Gaps 0;  
Best Local 67; Conservative 3.2%; 0; Mismatches 59; Indels 0;  
Matches  
Query Match similarity 3.4%; Score 31.6; DB 4; Length 4668; Gaps 0;  
Best Local 67; Conservative 3.2%; 0; Mismatches 59; Indels 0;  
Matches  
DB 378 gattcagacatg...  
DB 4525 gacttcgagatg...  
DB 438 ttctcactac...  
DB 4585 ttgtctatcac...  
DB 498 ttatc 503  
DB 4645 ttatcc 4650

RESULT 5 Application US/09195868  
US-09-195-868-11/C  
Sequence 11, Application US/09195868  
Patent No. 6090621  
GENERAL INFORMATION:  
APPLICANT: KAT P.H.D., DAVID  
INVENTOR: WILLIAMS M.D., DAVID  
ATTORNEY: WILLIAMS M.D., DAVID  
TITLE OF INVENTION: 33  
TITLE OF SEQUENCE: 33  
NUMBER OF SEQUENCES: 33  
CORRESPONDENCE: CHIRON CORPORATION  
ADDRESS: 4560 HORTON STREET  
STREET: EMERYVILLE  
CITY: EMERYVILLE  
STATE: CA  
COUNTRY: USA  
ZIP: 94608  
COMPUTER READABLE FORM: disk  
MEDIUM TYPE: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0; Version #1.30  
CURRENT APPLICATION NUMBER: US/09/195,868  
FILING DATE: 1997-03-21  
CLASSIFICATION NUMBER: US/09/195,868  
ATTORNEY/AGENT INFORMATION:

Wed Jan 9 14:38:17 2002

NAME: FIRESTONE, LEIGH H.  
REGISTRATION NUMBER: 36, 831  
TELECOMMUNICATION INFORMATION: 1182, 004  
TELEPHONE: 510-923-2707  
TELEFAX: 510-655-3470  
SEQUENCE FOR SEQ ID NO: 2  
LENGTH: 2940 base pairs  
TYPE: nucleic acid pairs  
STRANDEDNESS: single  
MOLECULE TYPE: linear  
US-09-195-868-11 CDNA

Query Match  
Best Local  
Matches 45; Similarity 3.38; Score 31.2; DB 3; Length 2940;  
OY 549 catctgtctaacgtgtgacccgcatcaaaatgtctgtcgatctctctcttcttggc 608  
Db 1641 CAAGCAGCTCGAGTGAACCTCAGTGAATTTGGTCTGGACTTGGCTTCAATGTGCG 1582  
OY 609 atgcacg 616  
Db 1581 ATAGCAGC 1574

RESULT 6  
US-09-195-868-12/C  
Sequence 12, Application US/09195868  
GENERAL No. 6090621  
APPLICANT: KAYNAUGH MD, MICHAEL  
APPLICANT: FOR PH.D., DAVID  
TITLE OF INVENTION: WILLIAMS MDPH, LEWIS T.  
NUMBER OF INVENTION: SIGNALING INOSITOL POLYPHOSPHATE  
CORRESPONDENCE ADDRESSES: 33  
STREET: CHIRON CORPORATION  
CITY: EMERYVILLE  
STATE: 4560 HORTON STREET  
COUNTRY: CA  
ZIP: 94608  
COMPUTER READABLE FORM:  
OPERATING SYSTEM: IBM PC compatible  
CURRENT APPLICATION DATA: PC-DOS/MS-DOS  
APPLICATION DATA: Release #1.0, Version #1.30  
CLASSIFICATION:  
FILING DATE: US/09/195, 868  
ATTORNEY/AGENT INFORMATION:  
NAME: FIRESTONE, LEIGH H.  
REGISTRATION NUMBER: 36, 831  
TELECOMMUNICATION INFORMATION: 1182, 004  
TELEPHONE: 510-923-2707  
TELEFAX: 510-655-3470  
SEQUENCE FOR SEQ ID NO: 2  
LENGTH: 2940 base pairs  
TYPE: nucleic acid pairs  
STRANDEDNESS: single  
MOLECULE TYPE: linear  
US-09-195-868-12 CDNA

US-09-740-211-13-copy\_11000\_11933.rni

Query Match  
Best Local  
Matches 45; Similarity 3.38; Score 31.2; DB 3; Length 3691;  
OY 549 catctgtctaacgtgtgacccgcatcaaaatgtctgtcgatctctcttcttggc 608  
Db 2392 CAGCAGCTCGAGTGAACCTCAGTGAATTTGGTCTGGACTTGGCTTCAATGTGCG 2333  
OY 609 atgcacg 616  
Db 2332 ATAGCAGC 2325

RESULT 7  
US-08-560-005-1/C  
Sequence 1, Application US/08560005  
GENERAL No. 6001354  
APPLICANT: ROBERT  
APPLICANT: WILLIAMS, DAVID A.  
APPLICANT: JEFFERSON, ANNE BENNETT  
TITLE OF INVENTION: Acids Encoding Therefor  
NUMBER OF INVENTION: No. 6001354el Gb2 Associating Protein and Nucleic  
CORRESPONDENCE ADDRESSES: 10  
STREET: One Mansend and Townsend and Crew  
CITY: San Francisco  
STATE: California  
COUNTRY: California  
ZIP: 94105  
COMPUTER READABLE FORM:  
OPERATING SYSTEM: IBM PC compatible  
CURRENT APPLICATION DATA: PC-DOS/MS-DOS  
APPLICATION DATA: Release #1.0, Version #1.25  
CLASSIFICATION:  
FILING DATE: US/08/560, 005  
ATTORNEY/AGENT INFORMATION:  
NAME: DOW, KAREN B.  
REGISTRATION NUMBER: 29, 684  
TELECOMMUNICATION INFORMATION: 2307K-0624000  
TELEPHONE: 415-326-2422  
TELEFAX: 415-326-2422  
SEQUENCE FOR SEQ ID NO: 1  
LENGTH: 4147 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
MOLECULE TYPE: linear  
FEATURE: DNA (genomic)  
NAME/KEY: CDS  
LOCATION: 17..2944  
US-08-560-005-1

Query Match  
Best Local  
Matches 45; Similarity 3.38; Score 31.2; DB 3; Length 4147;  
OY 549 catctgtctaacgtgtgacccgcatcaaaatgtctgtcgatctctcttcttggc 608  
Db 1648 CAGCAGCTCGAGTGAACCTCAGTGAATTTGGTCTGGACTTGGCTTCAATGTGCG 1582  
OY 609 atgcacg 616  
Db 2332 ATAGCAGC 2325

Db 1588 ATAGCACC 1581

## RESULT 8

US-09-418-540-1/C  
Sequence 1, Application US/09418540  
Patent No. 6296848  
GENERAL INFORMATION:  
APPLICANT: Pot. David A.  
APPLICANT: Williams, Lewis T.  
APPLICANT: Jefferson, Anne Bennett  
APPLICANT: Majerus, Philip W.  
TITLE OF INVENTION: No. 6296848el Gb2 Associating Protein and Nucleic  
TITLE OF INVENTION: Actin Encoding Therefor  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESS: Townsend and Townsend and Crew  
STREET: One Market Plaza, Stuart Tower, Suite 2000  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/418,540  
FILING DATE: 14-OCT-1999  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/560,005  
FILING DATE: 17-NOV-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Dow, Karen B.  
REGISTRATION NUMBER: 29,684  
REFERENCE/DOCKET NUMBER: 2307K-0624000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-326-2422  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4147 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 17..2944  
US-09-418-540-1

Query Match 3.3%; Score 31.2; DB 4; Length 4147;  
Best Local Similarity 66.2%; Pred. No. 1.9;  
Matches 45; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Db 1588 ATAGCACC 1581

RESULT 9  
US-08-664-962B-7/C  
Sequence 7, Application US/08664962B  
Patent No. 6218162  
GENERAL INFORMATION:  
APPLICANT: Krystal, Gerald

## TITLE OF INVENTION: SH2-CONTAINING INOSITOL-PHOSPHATASE

NUMBER OF SEQUENCES: 24

CORRESPONDENCE ADDRESS:

ADDRESSEE: MERCHANT &amp; GOULD

STREET: 3100 No. 6218162west Center, 90 South Seventh Street

CITY: Minneapolis

STATE: Minnesota

COUNTRY: U.S.A.

ZIP: 55402-4131

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/664,962B

FILING DATE: 14-JUN-1996

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Mueller, Douglas P.

REGISTRATION NUMBER: 30,300

REFERENCE/DOCKET NUMBER: M&amp;G 7933.49-US-01

TELEPHONE: 612-332-5300

TELEFAX: 612-332-9081

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 4870 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

ORIGINAL SOURCE:

ORGANISM: Homo sapiens

IMMEDIATE SOURCE:

CLONE: hSHIP

FEATURE:

NAME/KEY: CDS

LOCATION: 113..3673

US-08-664-962B-7





RESULT 15

US-08-337-483-112/c

; Sequence 112, Application US/08337483  
; Patent No. 5922562

## GENERAL INFORMATION:

APPLICANT: Loosmore, Sheena  
 APPLICANT: Harkness, Roblin  
 APPLICANT: Schryvers, Anthony  
 APPLICANT: Chong, Pele  
 APPLICANT: Gray-Owen, Scott  
 APPLICANT: Yang, Yan-ping  
 APPLICANT: Mardin, Andrew  
 APPLICANT: Klein, Michel  
 TITLE OF INVENTION: Transferrin Receptor Genes  
 NUMBER OF SEQUENCES: 147  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Sim & McBurney  
 STREET: Suite 701, 330 University Avenue  
 CITY: Toronto  
 STATE: Ontario  
 COUNTRY: Canada  
 ZIP: M5G 1R7

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/337,483  
 FILING DATE: 08-NOV-1994  
 CLASSIFICATION: 405

## ATTORNEY/AGENT INFORMATION:

NAME: Stewart, Michael I  
 REGISTRATION NUMBER: 24,973  
 REFERENCE/DOCKET NUMBER: 1038-410 MIS:jb  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (416) 595-1155  
 TELEFAX: (416) 595-1163  
 INFORMATION FOR SEQ ID NO: 112:

## SEQUENCE CHARACTERISTICS:

LENGTH: 1951 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear

## FEATURE:

NAME/KEY: CDS  
 LOCATION: 1..1890  
 US-08-337-483-112

## Query Match

Best Local Similarity 3.3%; Score 30.8; DB 2; Length 1951;  
 Matches 59; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 777 gctctgtttgcaatataaccgacgagctctgtacatgcaatctctgcatctgccc 836  
 DB 985 gctgtctttggcacttaataacccaataaactttttatccctagctaaanaattttcccc 926  
 QY 837 ccgagcgtcgcgactacgacgaataatccgataagcgaatgttc 882  
 DB 925 ctatgtcttcacacatttagcccatataaaccacaccttctaatgtttcc 880

Search completed: January 8, 2002, 17:21:33  
 Job time: 12344 sec

GenCore version 4.5  
Copyright (c) 1993-2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 8, 2002, 15:13:16; Search time 4541.38 Seconds  
(without alignments)  
2210.024 Million cell updates/sec

Title: US-09-740-211-13\_COPY\_11000\_11933

Perfect score: 934  
Sequence: 1 aaagcgagcgaggtgtgtag.....cttaagctgcctgcag 934

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: EST:

- 1: em\_estfun:\*\*
- 2: em\_esthug:\*\*
- 3: em\_esthug:\*\*
- 4: em\_esthug:\*\*
- 5: em\_esthug:\*\*
- 6: em\_esthug:\*\*
- 7: em\_esthug:\*\*
- 8: em\_esthug:\*\*
- 9: em\_esthug:\*\*
- 10: em\_esthug:\*\*
- 11: em\_esthug:\*\*
- 12: em\_esthug:\*\*
- 13: em\_esthug:\*\*
- 14: em\_esthug:\*\*
- 15: em\_esthug:\*\*
- 16: em\_esthug:\*\*
- 17: em\_esthug:\*\*
- 18: em\_esthug:\*\*
- 19: em\_esthug:\*\*
- 20: em\_esthug:\*\*
- 21: em\_esthug:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	724.4	77.6	734	10	AV404408 AV404408
2	605	64.8	628	10	AV404408 ESTTRKA-8
3	588	63.0	589	11	BG451667 NF092G10D
4	569.6	61.0	656	10	AV404408
5	531.4	56.9	546	10	AV404408
6	522.2	55.9	714	10	AV404408
7	455.4	48.8	743	10	AV404408
8	454.2	48.6	500	10	AV404408
9	420.6	45.0	617	11	BF328036
10	419.4	44.9	685	11	BF634302
11	387.2	41.5	688	11	BF269141
12	380.2	40.7	775	10	AV404200

13	373	39.9	373	11	BG312470
14	371.8	39.8	498	10	AA841579
15	360	38.5	399	10	AV749428
16	358.2	36.2	376	10	BE434444
17	338.2	35.0	329	10	BE353471
18	316.2	33.9	644	11	B1268701
19	305.2	32.7	408	11	B1268709
20	298	31.9	305	10	AM619942
21	294.6	31.5	558	11	B1268787
22	294.2	31.5	606	11	B1268699
23	286.4	30.7	366	11	B1268418
24	283.4	30.3	501	10	BE458797
25	260.6	27.9	291	11	N84801
26	246	26.3	246	11	BG459413
27	226.6	24.3	248	11	B1268786
28	217.2	23.3	475	11	BE934675
29	215	23.0	215	11	BE934207
30	184.4	19.7	211	11	BG312975
31	175	18.7	175	10	AV404308
32	154.4	16.5	242	10	AI898370
33	128.2	13.7	479	10	AI490216
34	126.4	13.0	193	11	C22175
35	121.8	13.0	269	11	C22255
36	117.2	12.5	401	11	BF657675
37	110.4	11.8	231	11	BF190428
38	94	10.1	412	11	BF552927
39	87	9.3	121	10	AW038061
40	87	9.3	393	11	BG451510
41	83.2	8.9	518	11	BG102333
42	83.2	8.9	578	11	BG053849
43	74	7.9	906	11	CNSOLUTV
44	67	7.2	578	11	B1263664
45	61.2	6.6	654	11	B1269244

ALIGNMENTS

RESULT 1  
AV404408 734 bp mRNA EST 06-FEB-2000  
AV404408 Bombyx mori pheromone gland Shuko x Ryuhaku newly-eclosed  
adult Bombyx mori cDNA clone pg--0946 T3, mRNA sequence.

ACCESSION AV404408.1 GI:6908496  
VERSION AV404408  
KEYWORDS EST.  
SOURCE domestic silkworm.  
ORGANISM Bombyx mori  
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditylaria  
; Bombycoidea; Bombycidae; Bombyx.

REFERENCE 1 (bases 1 to 734)  
Mita, K., Morimoto, M., Shimada, T., Okano, K. and Maeda, S.  
Bombyx mori cDNA  
unpublished (2000)  
JOURNAL Contact: Mita K  
COMMENT Genome Research Group  
National Institute of Radiological Sciences  
Anagawa 4-9-1, Inage, Chiba 263-8555, Japan  
Email: kmita@nirs.go.jp  
method:uni-directional, sequence direction:sequenced from T3 primer  
(5' -> 3')

FEATURES  
source  
/organism="Bombyx mori"  
/strain="Shuko x Ryuhaku"  
/db\_xref="taxon:7091"  
/clone="pg--0946"  
/clone\_id="Bombyx mori pheromone gland Shuko x Ryuhaku  
newly-eclosed adult"

Project="Silkworm Genome Program in MAF, and Research for the  
Future Program in JSPS", see 'Silkbase',  
<http://www.ab.a.u-tokyo.ac.jp/silkbase/>, for whole ESTdb.  
Location/Qualifiers  
1..734

/sex="female"  
/tissue\_type="pheromone gland"  
/dev\_stage="newly-eclosed adult"  
BASE COUNT 173 a 200 c 160 g 201 t  
ORIGIN

Query Match 77.6%; Score 724.4; DB 10; Length 734;  
Best Local Similarity 99.2%; Pred. No. 3.7e-211;  
Matches 728; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 170 tccggtctgctccggtatcgcttaatacagacaggaagagagcgtgtaccca 229  
DB 1 tccgggtctgctccggtatcgcttaatacagacaggaagagagcgtgtaccca 60  
QY 230 ggcctgcgcagccttcacagattatcgatcaatgcttaccgtgtgtcagataaaa 289  
DB 61 ggctgcgcagccttcacagattatcgatcaatgcttaccgtgtgtcagataaaa 120  
QY 290 aatccgaacacgtttagcagagcttcaactatcactgtgcgaactgttcggagttgcat 349  
DB 121 aatccgaacacgtttagcagagcttcaactatcactgtgcgaactgttcggagttgcat 180  
QY 350 ttgcagacacctctgcctgcagatgtgtgagttcagacagatacgtcgaatgacccaact 409  
DB 181 ttgcagacacctctgcctgcagatgtgtgagttcagacagatacgtcgaatgacccaact 240  
QY 410 aggcgagatcgtgagtaagagcgcgcctcttcaactcactacccaagagagataaa 469  
DB 241 aggcgagatcgtgagtaagagcgcgcctcttcaactcactacccaagagagataaa 300  
QY 470 cccatcgtttagcagacacattacccaatttattcaataagtaataatgacgttata 529  
DB 301 cccatcgtttagcagacacattacccaatttattcaataagtaataatgacgttata 360  
QY 530 tttgtgacacacgttgcagacacacacacacacacacacacacacacacacacacac 589  
DB 361 tttgtgacacacgttgcagacacacacacacacacacacacacacacacacacacac 420  
QY 590 attgacacctctgtgtgac 649  
DB 421 attgacacctctgtgtgac 480  
QY 650 gaccaggtgagttgagtaagtttggagttgagttgagttgagttgagttgagttgag 709  
DB 481 gaccaggtgagttgagtaagtttggagttgagttgagttgagttgagttgagttgag 540  
QY 710 gctgacacctctgtgtgac 769  
DB 541 gctgacacctctgtgtgac 600  
QY 770 ccccaacacctctgtgtgac 829  
DB 601 ccccaacacctctgtgtgac 660  
QY 830 ctgtcccccgcgttcgcgcagccttcagcaataatcgcgaataagtggtgcgaact 889  
DB 661 ctgtcccccgcgttcgcgcagccttcagcaataatcgcgaataagtggtgcgaact 720  
QY 890 tgcagttacattgc 903  
DB 721 tgcagttacattgc 734

RESULT 2  
AI065168/c AI065168 628 bp mRNA EST 24-JUL-1998  
LOCUS ESTTRKA-8 Rat Lambda Zap II Library, StrataGene #36501 Rattus  
DEFINITION norvegicus cDNA clone pGEM-T/TRKA-8MF 3, mRNA sequence.  
ACCESSION AI065168.1 GI:3340575  
VERSION EST.  
KEYWORDS Norway rat.  
SOURCE

## ORGANISM

Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

## REFERENCE

1 (bases 1 to 628)  
Bingdong, L. and Huang, B.R.  
Research on mechanism of p75NTR induced apoptosis (Bingdong, L. and  
Huang, B.R.)  
Unpublished (1998)  
Contact: Dongbing Lai  
National Laboratory of Medical Molecular Biology  
Chinese Academy of Medical Sciences & PUMC  
5 Dongdan Sanliao, Beijing 100005, P.R.China  
Tel: 86-10-65296406  
Email: huangbr@cdm.jimcoms.ac.cn  
Seq primer: M13 Forward Primer.  
Location/Qualifiers  
1..628

## JOURNAL

FEATURES  
source  
/organism="Rattus norvegicus"  
/strain="Sprague-Dawley"  
/db\_xref="taxon:10116"  
/clone="pGEM-T/TRKA-8MF"  
/clone\_1ib="Rat Lambda Zap II Library, StrataGene #36501"  
/sex="Male"  
/tissue\_type="Brain"  
/dev\_stage="6 weeks"  
/lab\_host="XLI-Blue MRP"  
/note="Vector: Lambda Zap II; Site\_1: EcoR I; Site\_2: Xho  
I"

BASE COUNT 173 a 138 c 166 g 148 t 3 others  
ORIGIN

Query Match 64.8%; Score 605; DB 10; Length 628;  
Best Local Similarity 98.7%; Pred. No. 1.5e-174;

Matches 608; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 192 gcttaattagacacagaaagacagcgttgcagacagcgttgcagacagcgttgcagac 251  
DB 628 gcttaattagacacagaaagacagcgttgcagacagcgttgcagacagcgttgcagac 569  
QY 252 atcgaactaagctcttgcagacagcgttgcagacagcgttgcagacagcgttgcagac 311  
DB 568 atcgaactaagctcttgcagacagcgttgcagacagcgttgcagacagcgttgcagac 509  
QY 312 ctctactattacactgcagacagcgttgcagacagcgttgcagacagcgttgcagac 371  
DB 508 ctctactattacactgcagacagcgttgcagacagcgttgcagacagcgttgcagac 449  
QY 372 tgggtgaggttcacagacagcgttgcagacagcgttgcagacagcgttgcagac 431  
DB 448 tgggtgaggttcacagacagcgttgcagacagcgttgcagacagcgttgcagac 389  
QY 432 cgcctcttcatctcactacacacacacacacacacacacacacacacacacacac 491  
DB 388 cgcctcttcatctcactacacacacacacacacacacacacacacacacacacac 329  
QY 492 ccaatttattcaataagcattatcagcgttgaatagtgccatcgcgttggaactat 551  
DB 328 ccaatttattcaataagcattatcagcgttgaatagtgccatcgcgttggaactat 269  
QY 552 gctgtaagttgagacagcattcaaaatgtgtcgtgagttgagttgagttgagttgag 611  
DB 268 gctgtaagttgagacagcattcaaaatgtgtcgtgagttgagttgagttgagttgag 209  
QY 612 gcaaccacagagcgttcaacagcgttcaacagcgttcaacagcgttcaacagcgttcaac 671  
DB 208 gcaaccacagagcgttcaacagcgttcaacagcgttcaacagcgttcaacagcgttcaac 149  
QY 672 ttggagttacatcgttcacagcgttcaacagcgttcaacagcgttcaacagcgttcaac 731  
DB 148 ttggagttacatcgttcacagcgttcaacagcgttcaacagcgttcaacagcgttcaac 89



OY	732	cgccttgcagcttcgcgacactcctttccagacaactotcccccaagctctgtttggcaa	791
Db	88	CGCCTTGCACTTCGCGACTGTCTGCACAACTCCGCCACAGCTCTGTTGGCAA	29
OY	792	tatcaaccgacaggcc	807
Db	28	TATCAACAGCACGCGCC	13
RESULT	3		
BG451667		589 bp	mRNA EST
LOCUS			16-MAR-2001
DEFINITION	NF092G1DPT1P1081 Drought Medicago truncatula cDNA clone NF092G1DPT		
ACCESSION	BG451667		
VERSION	5, mRNA sequence.		
KEYWORDS	BG451667.1 GI:13370461		
SOURCE	EST.		
ORGANISM	barrel medic. Medicago truncatula Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosoids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.		
REFERENCE	1 (bases 1 to 589) Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J., Flores, H.R., Imman, J.T., Weller, J.W. and May, G.D. Expressed Sequence Tags from the Samuel Roberts Noble Foundation Medicago truncatula drought library Unpublished (2000)		
JOURNAL	Contact: May GD Plant Biology Division The Samuel Roberts Noble Foundation 2510 Sam Noble Parkway, Ardmore, OK 73402, USA Tel: 580 221 7391 Fax: 580 221 7380 Email: gdmay@noble.org Insert Length: 589 Std Error: 0.00 Plate: 092 row: G column: 10 Seq primer: TCACACAGGAACCACTTATGAC.		
COMMENT	Location/Qualifiers 1..589 /organism="Medicago truncatula" /db_xref="taxon:3880" /clone_1fb="NF092G1DPF" /clone_1lb="Drought" /tissue_type="plantlets" /dev_stage="pooled timepoints" /note="Vector: Lambda Zap; Contains a mixture of entire plantlets harvested in a series of days-post-watering timepoints."		
FEATURES			
source	138 a 143 c 143 g 165 t		
BASE COUNT	138 a 143 c 143 g 165 t		
ORIGIN			
Query Match	63.0%; Score 588; DB 11; Length 589;		
Best Local Similarity	100.0%; Pred. No. 2.4e-169; Indels 0; Gaps 0.		
Matches	588; Conservative 0; Mismatches 0;		
OY	89	aattcagaacagacagtgcgtcaagctcagtttggtgtgtgtcattgtcgtcgagcgagatg	148
Db	2	AATTGAGACAGACAGTGGCTACGAGCTCAGTTCAGTTGGTTTGCTTCTTGCGGGCGGATG	61
OY	149	acgcctgtacgcatcttgatcgacggcttcgtcttcggatlcogcttaattcagcacac	208
Db	62	ACGCCTGTACGATTTGGTGGATCCGGTCTCTCCGATATTCCTTAATTACAGCACAC	121
OY	209	ggaagaagacatgcgtcaaacagagctcgagcgagctcttcacagatlaagactcaatgcctt	268
Db	122	GGAAGAAGCATCGGTAAACCAAGCTCGCGAGCTCTTCACAGATTATCGACTCAATGCTCTT	181
OY	269	acctgttgtgcagataaaaaaatcccgaqaacgltatgcagtgctctaactatactctgc	328
Db	182	ACCTGTTGTGCAGATTATAAAAAATCCCGAACCCGTTATGACAGGCTCTAATATTACCTGC	241

QY	329	gaactgttcaggattgcatttttcagaccctctctgtcctgcgaatggttgagatctccagac	388
Db	242	GAACCTGTTTGGGATTGCAATTTTGCAGACCTCTCTGCTCCGTCGATGTTGGAGTTCCAGAC	301
QY	389	gatacgttcgaagtgcaccacttcaggagaaatgcgtatgaagcagccgcctcttcatctca	448
Db	302	GATPACGTGCAAGTACCACCACTAGGCGGAATTCGGTACTAGTAAAGCCGCCCTCTTTCATCTCA	361
QY	449	ctaccacaacagcgaattaacccatctgtttgagtcacaaattacccaatttattcaataa	508
Db	362	CTACCAACAGCAGGCAATTAACCATCTGTATGATCAAAATTTCCCAATTTTATTCATTA	421
QY	509	gtcaatatcatgcgcgttaataatgtttgccatcgcgtgacaatcatctgtctcaacgttgacc	568
Db	422	GTCATATATCAATGCGCTTAATATGTTGGCATCTGTCGCAATTCATGCTGCTAACGTGTGAC	481
QY	569	gcattcaaatatgtctcgtcgcgattgactcttcttctgtgcatctgacccaccagagcgtca	628
Db	482	GCATTCAAAATGTTGCTCGCATTCATCTCTTCTTGGGATTCACACCAAGAGCGTCA	541
QY	629	taccagcgcttaacacagtcgctgaccacagtgatggttggttaagatgttggg	676
Db	542	TACACGCGCTTAACAGTGCATGACAGGTGGTGGTGGTAAAGTTTGGG	589
RESULT	4		
LOCUS	A1489217	656 bp	mRNA
DEFINITION	EST747556 tomato ovary, TMU Lycopersicon esculentum cDNA clone	18-MAY-2001	
ACCESSION	A1489217		
VERSION	A1489217		
KEYWORDS	EST.		
SOURCE	ORGANISM	tomato.	
		Lycopersicon esculentum	
		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
		Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;	
		Asteridae; euasterids I; Solanales; Solanaceae; Solanum;	
REFERENCE	1 (bases 1 to 656)		
AUTHORS	Alcala,J., Vrebalov,J., White,R., Matern,A.L., Vision,F., Holt,I.E.,		
	Liang,F., Upton,J., Rongning,C.M., Craven,M.B., Fujii,C.Y., Bowman		
	,C.L., Nierman,W., Fraser,C.M., Venter,J.C., Martin,G.B., Tanksley		
	,S.D. and Giovannoni,J.		
	Generation of ESTs from tomato carpel tissue		
TITLE	Unpublished (1999)		
JOURNAL	Contact: CUCI		
COMMENT	Clemson University Genomics Institute		
	Clemson University		
	100 Jordan Hall, Clemson, SC 29634, USA		
	Email: <a href="http://www.genome.clemson.edu/orders/index.html">http://www.genome.clemson.edu/orders/index.html</a> .		
FEATURES	Location/Qualifiers		
SOURCE	1..656		
	/organism="Lycopersicon esculentum"		
	/cultivar="TA96"		
	/db_xref="taxon:4081"		
	/clone="CLEDI7L7"		
	/clone_id="tomato ovary, TMU"		
	/tissue_type="carpel"		
	/dev_stage="5 days pre-anthesis to 5 days post-anthesis"		
	/lab_host="XL1-Blue MR"		
	/note="Vector: Bluescript SK(-); Site_1: EcoRI, Site_2:		
	XhoI; CLEP - Tomato Carpel EST library. OligodT-primed and		
	directionally cloned cDNA in vector lambda ZAP II with 5'		
	and 3' ends located at the EcoRI and XhoI sites,		
	respectively."		
BASE COUNT	166 a 159 c 153 g 176 t 2 others		
ORIGIN			
Query Match	61.0%;	Score 569.6;	DB 10;
Best Local Similarity	99.0%;	Pred. No. 1.1e-163;	Length 656;

Matches 572: Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 aaagcgagaggtgtgttagcgagcactcctgcacccgcttcaagaagtatgtgta 60  
 DB 79 AAAGCGAGAGGTTGTTAGCGCACTCCGACCCGCTTACGAAAGCATGTGTA 138  
 OY 61 aaagcgagaggtgtgttagcgagcactcctgcacccgcttcaagaagtatgtgta 120  
 DB 139 AAAGCGAGAGGTTGTTAGCGCACTCCGACCCGCTTACGAAAGCATGTGTA 198  
 OY 121 tgggt 180  
 DB 199 TGGGT 258  
 OY 181 ttccggtatcgcgttaattcaagacaaagaaagacactggtcaacagctgcgcac 240  
 DB 259 TTCCGGTATTCGCTTAATTCAGACAAAGAAAGACACTGCTAACCAGCTCCGAC 318  
 OY 241 tctcaagatcgcgttaattcaagacaaagaaagacactggtcaacagctgcgcac 300  
 DB 319 TCTTCAAGATTCGCACTCAATGCTCTTACGCTGTGTGTGTGTGTGTGTGTGTGTGT 378  
 OY 301 cgttagcagagctcaactaactaactgctgcaactgcttgcgagatgcatlgtgcagac 360  
 DB 379 CGTTATGACAGGCTTAACCTAATACCTGCAACTGTTTCCGATTTTGCAGACTTC 438  
 OY 361 tctgctgcatgt 420  
 DB 439 TCTGCTGCTGATGT 498  
 OY 421 gtagtaagcgagcctcttcttcaactcaactcaactcaactcaactcaactcaactca 480  
 DB 499 GTAGTAAGCGCGCTTCTTTCATCTCACTCACTCACTCACTCACTCACTCACTCACT 558  
 OY 481 gtaaatatcccaatttatttcaataatgaatcaatcaatcaatcaatcaatcaatcaat 540  
 DB 559 GTCAATTTTACCAATTTTATTCATAATGATCATGATGATGATGATGATGATGATGAT 618  
 OY 541 gtggaatcagctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 578  
 DB 619 GTGGCAATCATGCTGCTAATGCTGATGACCGCATTTAANA 656

RESULT 5  
 AM624097/c 546 bp mRNA EST 18-MAY-2001  
 LOCUS EST322042 tomato flower buds 3-8 mm, Cornell University  
 DEFINITION Lycopersicon esculentum cDNA clone cTOB14J1 5', mRNA sequence.  
 ACCESSION AM624097  
 VERSION AM624097.1 GI:7337124  
 KEYWORDS EST.  
 SOURCE tomato.  
 ORGANISM Lycopersicon esculentum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Asteridae; euasterids I; Solanales; Solanaceae; Solanum;  
 Lycopersicon  
 1 (bases 1 to 546)  
 van der Hoeven, R.S., Bezzerides, J.L., Matern, A.L., Holt, I.E., Liang,  
 F., Hansen, T., Craven, M.B., Bowman, C.L., Rönning, C.M., Niernau, W.,  
 Fraser, C.M., Martin, G.B., Giovannoni, J.J. and Tanksley, S.D.  
 Generation of ESTs from tomato flower tissue, 3-8 mm buds  
 Unpublished (1999)  
 CONTACT: CUGI  
 CLEMSON UNIVERSITY GENOMICS INSTITUTE  
 CLEMSON UNIVERSITY  
 100 JORDAN HALL, CLEMSON, SC 29634, USA  
 EMAIL: <http://www.genome.clemson.edu/orders/index.html>  
 5 prime sequence:  
 Location:Qualifiers  
 1. 546  
 /organism="Lycopersicon esculentum"  
 /cultivar="TA96"

## FEATURES

source

/db\_xref="taxon:4081"  
 /clone="cTOB14J1"  
 /clone\_1b="tomato flower buds 3-8 mm, Cornell University"  
 /lssue\_type="flower"  
 /dev\_stage="3-8mm buds"  
 /note="Vector: pBluescript SK(-); Site\_1: EcorI; Site\_2:  
 taken from greenhouse plants (4-8 wks old, TA96). They  
 were immediately frozen in liquid nitrogen and then  
 size-separated while remaining frozen."  
 BASE COUNT 157 a 124 c 135 g 130 t  
 ORIGIN

Query Match 56.9%; Score 531.4; DB 10; Length 546;  
 Best Local Similarity 99.6%; Pred No. 5.4e-152;  
 Matches 543; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

OY 65 gccgcagcgttaactatctactaattgaattcaagacagacagatggtcagctcagttggg 124  
 DB 545 GCCGCAGCGTAACTAATTAATTCAGACAGACAGACAGATGGCTACGCTCATTTGGC 486  
 OY 125 ttgtctgtctgt 184  
 DB 485 TTGTGCTGT 426  
 OY 185 ggtatcgtcttaattcaagacagacagacagacagacagacagacagacagacagac 244  
 DB 425 GGTATTCGCTTAATTCAGACAGACAGACAGACAGACAGACAGACAGACAGACAGAC 366  
 OY 245 caagattatcgcgttaactaactgcttcaactgctgtgtgcagatataaaaaaaccgaa 304  
 DB 365 CACGATTTATGACATTAATGCTTTACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 306  
 OY 305 atgcagcgttaactaactaactgcaactgcttgcgagatgcatlgtgcagacactctg 364  
 DB 305 ATGCAGGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 246  
 OY 365 cctgcagatgt 424  
 DB 245 CCTGCGATGT 186  
 OY 425 taagcgccgctcttcttcaactcaactcaactcaactcaactcaactcaactcaactca 484  
 DB 185 TAAGCGCGGCTCTTTCATCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCA 126  
 OY 485 aattaccgaatttatttcaataatgaatcaatcaatcaatcaatcaatcaatcaatcaat 544  
 DB 125 AATTACCGAATTTTATTCATAATGATCATGATGATGATGATGATGATGATGATGATGAT 66  
 OY 545 caatcagctgctaac-gtctgacccattcaaaaagtgtctgcatlgtgtctcttctt 603  
 DB 65 CAATCATGCTGTGTAACGGGTGTGACCGCATTTCAAAATGTTGTGTGATGATGATGATGAT 6  
 OY 604 gtggc 608  
 DB 5 GTGGC 1

RESULT 6  
 AV404019 714 bp mRNA EST 06-FEB-2000  
 LOCUS AV404019 Bombyx mori pheromone gland Shuko x Ryuhaku newly-eclosed  
 DEFINITION adult Bombyx mori cDNA clone pg--0345 T3, mRNA sequence.  
 ACCESSION AV404019  
 VERSION AV404019.1 GI:6908107  
 KEYWORDS EST.  
 SOURCE domestic silkworm.  
 ORGANISM Bombyx mori  
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia  
 ; Bombycoidea; Bombycidae; Bombyx.  
 1 (bases 1 to 714)

AUTHORS Mita, K., Morimyo, M., Shimada, T., Okano, K. and Maeda, S.  
 TITLE Bombix mori cDNA  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Mita K  
 Genome Research Group  
 National Institute of Radiological Sciences  
 Anagawa 4-9-1, Inage, Chiba 263-8555, Japan  
 Email: kmitsa@nirs.go.jp  
 method: uni-directional, Sequence direction: sequenced from T3 primer  
 (5' -> 3')  
 Project="Silkorm.Genome Program in MAFF, and Research for the  
 Future Program in JSPS", see "Silkbase"  
 Future Program in JSPS, see "Silkbase"/>, for whole ESTdb.  
 Location/Qualifiers

## FEATURES

Source  
 1. 714  
 /organism="Bombix mori"  
 /strain="Shuko x Ryuhaku"  
 /db\_xref="taxon:7091"  
 /clone\_lib="Bombix mori pheromone gland Shuko x Ryuhaku  
 newly-eclosed adult"  
 /sex="female"  
 /tissue\_type="pheromone gland"  
 /dev\_stage="newly-eclosed adult"  
 BASE COUNT 160 a 200 c 158 g 195 t 1 others  
 ORIGIN

Query Match 55.9%; Score 522.2; DB 10; Length 714;  
 Best Local Similarity 99.2%; Pred. No. 4e-149;  
 Matches 524; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
 401 TGACCAACTAGCGGAGTACGTAAGTACCGGCTCTTTTCATCTCATTACCAACGA 60  
 1 TGACCAACTAGCGGAGTACGTAAGTACCGGCTCTTTTCATCTCATTACCAACGA 60  
 461 GCGAATTAACCCATCGTGTAGTCAATTTACCAATTTATCAATGATCAATATCATG 120  
 61 GCGAATTAACCCATCGTGTAGTCAATTTACCAATTTATCAATGATCAATATCATG 120  
 521 CCGTAAATATGTGCGATCCGTCGCAATATCTCTAATCGTACCGCATTCATAATG 180  
 121 CCGTAAATATGTGCGATCCGTCGCAATATCTCTAATCGTACCGCATTCATAATG 180  
 581 TTGCTGCGATGACTTCTTGTGCGATGACACACGCGTCAATACGCGCTTA 240  
 181 TTGCTGCGATGACTTCTTGTGCGATGACACACGCGTCAATACGCGCTTA 240  
 641 ACAAGTGCAGTACGAGTGGGTGGTAAAGTTGGATTAAGCATGCTACACGCGATAT 300  
 241 ACAAGTGCAGTACGAGTGGGTGGTAAAGTTGGATTAAGCATGCTACACGCGATAT 300  
 701 GCTGCGCTGCTGCGATCTTGAATACCGACGCTTTGATCTTCGACACTTCTCG 360  
 301 GCTGCGCTGCTGCGATCTTGAATACCGACGCTTTGATCTTCGACACTTCTCG 360  
 761 ACAACTCTCCACACAGCTCTGTTTGGCAATATCAACCGCAGGCTGACCATGCA 420  
 361 ACAACTCTCCACACAGCTCTGTTTGGCAATATCAACCGCAGGCTGACCATGCA 420  
 821 TCTCTGCACTTGTGCGCGGCGTCCGCGCATACGCAATAATCCGATTAAGGCAATGTT 480  
 421 TCTCTGCACTTGTGCGCGGCGTCCGCGCATACGCAATAATCCGATTAAGGCAATGTT 480  
 881 GCGAGCACTTGCAGTACCTTGTGCTTGAATTTCTTCAACCTTTGCC 528  
 481 GCGAGCACTTGCAGTACCTTGTGCTTGAATTTCTTCAACCTTTGCC 528

RESULT 7 BE420137 743 bp mRNA EST 24-JUL-2000  
 LOCUS BE420137 743 bp mRNA EST 24-JUL-2000  
 DEFINITION WMS020.E9R000101.1 ITBC WMS wheat Scutellum Library Trilicium aestivum

ACCESSION BE420137  
 VERSION BE420137.1 GI:9417983  
 KEYWORDS EST.  
 SOURCE bread wheat.  
 ORGANISM Trilicium aestivum

REFERENCE  
 Anderson, O.A., Appels, R., Bailey, P., Blake, T., Close, T., Cloutier  
 S., Dubcovsky, J., Feuillet, C., Gale, M., Graner, A., Gustafson, P.,  
 Herrmann, R.G., Holtan, T., Jacquemyn, J.M., Jia, J., Joudrier, P.,  
 Langridge, P., Lazo, G.R., Lin, J.J., McGuire, P., Ogihara, T.,  
 Pecchioni, N., Quailset, C., Schuch, W., Selvaraj, G., Shariflou, M.,  
 Sorrells, M., Warburton, M., and Wenzel, G.  
 International Trilicium EST Cooperative (ITEC): Production of  
 Expressed Sequence Tags for Species of the Trilicaceae  
 Unpublished (2000)  
 Contact: Schuch W

JOURNAL  
 COMMENT Zeneca Wheat Improvement Centre, Norwich Research Park  
 Colney Lane, Norwich NR4 7UH UNITED KINGDOM  
 Tel: 44 1603 250 2600  
 Fax: 44 1603 250 699  
 Email: wolfgang.schuch@zeneca.com  
 International Trilicium EST Cooperative (ITEC)  
 http://wheat.pw.usda.gov/genome.

## FEATURES

Source  
 1. 743  
 /organism="Trilicium aestivum"  
 /cultivar="Novosibirskaya 67"  
 /db\_xref="taxon:4565"  
 /clone="WMS020.E9"  
 /clone\_lib="ITBC WMS wheat Scutellum Library"  
 /tissue\_type="scutellum callus"  
 /note="M13 Reverse sequencing primer used for 5' end of  
 clone."

BASE COUNT 148 a 231 c 212 g 152 t  
 ORIGIN

Query Match 48.8%; Score 455.4; DB 10; Length 743;  
 Best Local Similarity 99.8%; Pred. No. 1.3e-128;  
 Matches 456; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 397 GAAGTACCACTAGCGGAGTGGTAAAGTACCGGCTCTTTTCATCTCATTACCA 456  
 1 GAAGTACCACTAGCGGAGTGGTAAAGTACCGGCTCTTTTCATCTCATTACCA 456  
 457 ACAAGTGCAGTACGAGTGGGTGGTAAAGTTGGATTAAGCATGCTACACGCGATAT 516  
 61 ACAAGTGCAGTACGAGTGGGTGGTAAAGTTGGATTAAGCATGCTACACGCGATAT 516  
 517 CAGCGCTAATATGTTGCGATCCGTCGCAATCAATGCTGCTAAGTGTGACCGCATTC 180  
 121 CAGCGCTAATATGTTGCGATCCGTCGCAATCAATGCTGCTAAGTGTGACCGCATTC 180  
 577 AATGCTGCTGCGATGACTTCTTGTGCGATGACACACGAGCGTCAATACGCG 240  
 181 AATGCTGCTGCGATGACTTCTTGTGCGATGACACACGAGCGTCAATACGCG 240  
 637 CTAACAGTGCAGTACGAGTGGTGGTAAAGTTGGATTAAGCATGCTACACGCG 636  
 241 CTAACAGTGCAGTACGAGTGGTGGTAAAGTTGGATTAAGCATGCTACACGCG 636  
 697 ATATGCTGCGCTGCTGCGATCCGTCGCAATCAATGCTGCTAAGTGTGACCGCATTC 360  
 301 ATATGCTGCGCTGCTGCGATCCGTCGCAATCAATGCTGCTAAGTGTGACCGCATTC 360  
 757 CTCGACCACTCTCCCGACAGCTCTGTTTGGCAATATCAACCGCAGCGCTGTACATG 816  
 361 CTCGACCACTCTCCCGACAGCTCTGTTTGGCAATATCAACCGCAGCGCTGTACATG 816



QY	163 a	164 c	169 g	189 t
61	aaagcgcgca	-gcgt	aaactatt	aaatgaattcaagacagacagcgtacgctcagt 119
119				
442	AAAGGCCGCGAGCG	GAACATTACT	AAAGAAATTCAGSACAGACAGTGCCTACGGCTCACT	383
120	ttgggttcgtgcgtttg	ctgctg	ggcgagatgaagcgtctgacgcatlgtgltacggtlctg	179
179				
382	TTGGGTGTGTGCTGT	CTCTCTGCTGCGGAGATGAGCGCTGTACCATTTGGTATCGGATCTG		323
180	cttcggatctgcgttaatt	cagccagcaaggaagacatcgtctaccaggtctgcgca		239
239				
322	CTTCGGGTATTCGCT	TAATTCAGCAACAACGAAAGACACTGCTTAACACGGCTGCACGA		263
240	ctcttcacgatatcgcata	gtctctacccitctgtgcagatataaaatcccgaa		299
299				
262	CTCTTCACGATCATC	ATGCACATACGCTC-TACCTGTGTGCAGATATATAAAATCCGAAA		204
300	ccgtatgcagcgtcctaact	attaccctgcgcgaactglttcggatctgcatlgttcgagact		359
359				
203	CCGTTATGACGAGCT	CTACATTACTCTCGAAACGTTCGCGGATTCGATTTTCACACCT		144
360	ctctgcctgcgagatgg-	ttggaattccaaagcatatcgtcgaagtaccacgaagcgaat		418
418				
143	CTCGCTCGGATGATGAT	TGGAGTTCCAGACGATACGTGGAAGTACCACTAGGCGGAAT		84
419	cgtgagtaagcgcgcgcct	cttttcattcatctaccacgaag 461		
461				
83	CGTAGTAGAGCGCGC	CTCTTTTCATCTCATCTACACACAGCGCG 41		
RESULT 10				
BF634302	685 bp	mRNA	EST	19-DEC-2000
LOCUS	NE074F11D1F1094	Drought	Medicago truncatula	CDNA clone NE074F11DT
DEFINITION	5', mRNA sequence, 1			
ACCESSION	BF634302			
VERSION	BF634302.1	GI:11898460		
KEYWORDS	EST.			
SOURCE	barrel medic.			
ORGANISM	Medicago truncatula			
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;			
	Rosidae; eustosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae			
	Medicago.			
REFERENCE	1 (bases 1 to 685)			
AUTHORS	Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J., Flores, H.R., Imago, J.T., Meller, J.W. and May, G.D.			
TITLE	Expressed Sequence Tags from the Samuel Roberts Noble Foundation Medicago truncatula drought library			
JOURNAL	unpublished (2000)			
COMMENT	Contact: May GD Plant Biology Division, The Samuel Roberts Noble Foundation 2510 Sam Noble Parkway, Ardmore, OK 73402, USA Tel: 580 221 7391 Fax: 580 221 7380 Email: gdmay@noble.org Insert Length: 685 Std Error: 0.00 Plate: 074 row: F column: 11 Seq primer: TCACACAGAAACAGTATGAC.			
FEATURES	Location/Qualifiers			
source	1..685 /organism="Medicago truncatula" /db_xref="taxon:3880" /clone="NE074F11DT" /clone_11b="Drought" /tissue_type="Plantlets" /dev_stage="Pooled timepoints" /note="Vector: Lambda Zap; Contains a mixture of entire plantlets harvested in a series of days-post-watering timepoints."			
BASE COUNT	163 a	164 c	169 g	189 t
ORIGIN				

Query Match	44.9%	Score 419.4	DB 11	Length 685
Best Local Similarity	99.8%	Pred. No. 1,4e-117	Mismatches 1	Indels 0
Matches 420	Conservative	0	Mismatches 1	Gaps 0
OY	1	aaagcgagagaggtgttttagcggaactctcgccaccgcgcttcacgaagtcattgta 60		
Db	260	AAACGGCAGGAGTGTGTAGCGCACTCTCGCACCCGCTTTCAGCAAGTCATGTGTA 319		
OY	61	aaagccgcacgcgaactattactaatgaattccagaacagatggctacgcgtcagtt 120		
Db	320	AAAGGCCCCACGTAACATAATTAAATTCAGAGACAGACAGTGCCTACGGCTCAGTT 379		
OY	121	tggatgtgtcgttgctgtgctggcgcgatgacgcgcgttacacgatttggatccggttcgc 180		
Db	380	TGGGTTGGCTGTGTCCTGGCGCGATGACGCCCTGTACGCAATTGGTGATCCGGTTCCTGC 439		
OY	181	ttccgatttcgcttaattccagcaacaagagaagacacttgcttaaccaggcttcgcgac 240		
Db	440	TTCCGGTMTTGGCTTAATTACAGCAACAGGAAGAGCACTGGCTAACCCAGGCTCGCGAC 499		
OY	241	tcttcaagattatcgatcccaatagccttaccctgtgtgtgcagatataaaaaatcccgaaac 300		
Db	500	TCTTACGCTTATTCACGCTCAATGCTCTTACCTGTTGTGCAGATATATAAATTCGCGAATC 559		
OY	301	cgttatgcagcgtcccaactatcctctgcaacgttttcggagattgacatttgcagacctc 360		
Db	560	CGTTATGCAGGCTTAACTAATACCTGCGCACTGTTGGGATTCATTTGGCAACCTTC 619		
OY	361	tctgcctgcgatggtttgagatccagaacgatacgtccgaagtaccacttaggcgaatcg 420		
Db	620	TCGCTCGGATGTGGAGTTCACAGCATACGTCGAAGTACCACTAGGCGGAATCG 679		
OY	421	g 421		
Db	680	G 680		
RESULT 11				
BI269141	688 bp	mRNA	EST	18-JUN-2001
LOCUS	NE003A05IR1F1036	Irradiated Medicago truncatula cDNA clone		
DEFINITION	NE003A05IR 5', mRNA sequence.			
ACCESSION	BI269141			
VERSION	BI269141.1	GI:14875595		
KEYWORDS	EST.			
ORGANISM	Medicago truncatula			
SOURCE	barrel medic.			
REFERENCE	1 (bases 1 to 688)			
AUTHORS	Floris-Jerem, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J.,			
TITLE	Torres, H.R., Imman, J.T., Weller, J.W. and May, G.D.			
JOURNAL	Expressed Sequence Tags from the Samuel Roberts Noble Foundation			
COMMENT	Medicago truncatula irradiated library			
	Unpublished (2001)			
	Contact: May GD			
	Plant Biology Division			
	The Samuel Roberts Noble Foundation			
	2510 Sam Noble Parkway, Ardmore, OK 73402, USA			
	Tel: 580 221 7391			
	Fax: 580 221 7380			
	Email: gdmay@noble.org			
	Insert Length: 688	Std Error: 0.00		
	Plate: 003	row: A column: 05		
	Seq primer: TCACACGGAACAGCATATGAC.			
FEATURES	location/Qualifiers			
SOURCE	1..688			
	/organism="Medicago truncatula"			
	/db_xref="taxon:3880"			
	/clone="NE003A05IR"			

/clone\_1lb="Irradiated"  
/tissue\_type="seedling"  
/dev\_stage="seedling"  
/note="Vector: lambda Zap; Seedlings were exposed either to 100 Gy gamma or 0.5, 1, 5, or 10 kJ/m<sup>2</sup> UV irradiation. Gamma-irradiated samples were harvested at 6, 12, 24 and 48 hours after treatment. UV-irradiated samples were harvested 24 hours post-treatment. cDNA was prepared from poly(A<sup>+</sup> enriched, pooled samples of equivalent amounts of total RNA from each sample. The cDNA was directionally ligated into the Uni-Zap XR vector (Stratagene) and packaged using the Gigapack III Gold packaging extracts. Phagemids containing cDNA inserts were in vivo excised from the recombinant Uni-Zap XR vector using EXAssist helper phage and the E. coli strain XL1-Blue MRF' (Stratagene). Excised plasmids were plated using SOLR cells."

BASE COUNT 140 a 193 c 155 g 193 t 7 others  
ORIGIN

Query Match 41.5%; Score 387.2; DB 11; Length 688;  
Best Local Similarity 98.8%; Pred. No. 1e-107;  
Matches 411; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

OY 515 atcagccgttaatatgttgcacccgttgcaatcagc-tgtaacgttgaccgcatc 573  
DB 1 ATCATGCGCTTAATATGTCGCCATCGTCATCATGCTGCTAAACGTGACCGCAT 60  
OY 574 caaaatgtctgcagatgaactctctcttgatgacacccagagagcagcagc 633  
DB 61 CAAATGCTGTCGCGATTGACTCTTCTTGCGCTTACACCGACAGGCTCTATACAG 120  
OY 634 cagccttaacagtgctgagcagagtggttggaaggttgagatgaacatgctacagc 693  
DB 121 CGGCTTAACAGTCCGAGAGGAGTGGTGGTAAGGTTGGGATTAGCATGTCACAGC 180  
OY 694 ggcataatgc-tgagctgtctgcacatcttgatgaagcagccttgatcttcagcact 752  
DB 181 GCGATATGCTGGGCTTGGCTGGCATCTGATACCGACGCTTGTGATCTTCGCGACT 240  
OY 753 cttctcgcaaacctctcccaacagctctgttttgcaataatcaacagcagcagcgtgac 812  
DB 241 CTTTCTCGCAACTCTCCCGACAGCTCTGTTTGGCAATATCAACCGCGCTGTGAC 300  
OY 813 catggaactctgcatctgtgccccggcggtgcgagcctacgcagcaataatcgcagtaag 872  
DB 301 CATGGCAATCTCTGCACTTGGCCCCGGCGGTGCGGCACTACGGCAATATCCCGCATAG 360  
OY 873 cgaatgtcgagacacttgacacttgcttgatgaatttccttcaagctgcccc 928  
DB 361 CCAATGCTCGAGCACTTGCACTTGGCTTGAATGATTCTTCTTCAAGCTTTGCC 416

## RESULT 12

LOCUS AVA04200 775 bp mRNA EST 06-FEB-2000  
DEFINITION AVA04200 Bombyx mori pheromone gland Shuko x Ryuhaku newly-eclosed adult Bombyx mori cDNA clone pg-0693 T3, mRNA sequence.  
ACCESSION AVA04200  
VERSION AVA04200.1 GI:5908288  
KEYWORDS EST.  
SOURCE domestic silkworm.  
ORGANISM Bombyx mori

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea; Bombycidae; Bombyx.

REFERENCE 1 (bases 1 to 775)  
Mita, K., Morimoto, M., Shimada, T., Okano, K. and Maeda, S.

AUTHORS Bombyx mori cDNA  
TITLE Unpublished (2000)  
JOURNAL Contact: Mita K  
COMMENT Genome Research Group

National Institute of Radiological Sciences  
Anagawa 4-9-1, Inage, Chiba 263-8555, Japan  
Email: km1ta@nirs.go.jp  
Method: uni-directional, sequence direction: sequenced from T3 primer (5' -> 3')  
Project: 'Silkworm Genome Program in MAFF, and Research for the Future Program in JSPS'. see 'SilkBase',  
<http://www.ab.a.u-tokyo.ac.jp/silkbase/>, for whole ESTdb.  
Location/Qualifiers  
1. 775  
/organism="Bombyx mori"  
/strain="Shuko x Ryuhaku"  
/db\_xref="taxon:7091"  
/clone="pg-0693"  
/clone\_1lb="Bombyx mori pheromone gland Shuko x Ryuhaku newly-eclosed adult"  
/sex="female"  
/tissue\_type="pheromone gland"  
/dev\_stage="newly-eclosed adult"

BASE COUNT 161 a 211 c 185 g 218 t  
ORIGIN

Query Match 40.7%; Score 380.2; DB 10; Length 775;  
Best Local Similarity 99.2%; Pred. No. 1.5e-105;  
Matches 382; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 544 gcaatactgtctgaacgttgtagcagcattcaaaatgtgtctgcagatgaactctctt 603  
DB 1 GCAATATGCTGCTTAACGTTGACCGCATTTCAAAATGTTGCTGCGATTGACTCTTCT 60  
OY 604 gtgacattgaccacacagagcgctacacagcggtctaacagtggtgacaggtgggttg 663  
DB 61 GTGGCAATTGACACACAGACGCTCATACAGCGGCTTAACAGTGCCTGACAGGTGGTTG 120  
OY 664 ggtgaaggttgagatgacatgctacacagcgagatagctgctgcttgtagcacttga 723  
DB 121 GGTAAAGTTTGGGATTAGCATGTCACAGCGGAGATATCTGCGGTGGTCATCTTGA 180  
OY 724 atagccgagcctcttgacatctctccgacatcttctcgcgaacactctcccaacagctgt 783  
DB 181 ATAGCGGAGCGCTTTTGCACTTTCGCACTCTTTCGCAACTCTCCCGACACACTCTGT 240  
OY 784 ttggcaataataacacagcagcctgtacacatggcaatctctgacatctgccccggcg 843  
DB 241 TTTGGCAATATCAACCGACGCGCTGTACATGCAATCTTGTGATCTTGGCCCCGGCGT 300  
OY 844 cgcgagcactagcgcaataatccgataagcgaatgttgagacacttgcaatcttgc 903  
DB 301 CGGCGACTACGCGCAATATCCGATTAAGCGAATGTTGCGAGCACTTGCACTGACTTTCG 360  
OY 904 cctagatcttcctcaagctgcccc 928  
DB 361 CTTAGTATTTCCTTCAAGCTTTGCC 385

## RESULT 13

LOCUS BG312470 373 bp mRNA EST 23-FEB-2001  
DEFINITION WHE2408.F10.L20S Wheat 3-6 DAP seed cDNA library Triticum aestivum  
ACCESSION BG312470  
VERSION BG312470.1 GI:13114273  
KEYWORDS EST.  
SOURCE bread wheat.  
ORGANISM Triticum aestivum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae; Triticaceae; Triticum.

REFERENCE 1 (bases 1 to 373)  
Akhunov, E., Anderson, O.D., Chao, S., Chin, A., Choi, D.W., Close, T.J., Dvorak, J., Fenton, R.D., Gustafson, P., Han, P.S., Hsia, C.C., Kang, Y., Kianian, P., Lazo, G.R., Miller, R., Otto, C., Nguyen, H.T., Rausch, C.J.



```

Qy 276 gtgcagataaaatcccgaaacgctatgcagcgtcctaactatctacgtcgaactgt 335
Db 304 GTGCAGATATAAAATCCGAACCGTTATGCAGGCTCTACATATATACCTGGCAACTGT 363
Qy 336 ttgcgattgcatcttggaagacctctctgcctgcgtatggttggaagctcgaagatcgt 395
Db 364 TTCGGATTGGATTTGGACACCTCTGCTCGATGGTTGGAGTTCCACACGATAGT 423
Qy 396 cgaagtcacacacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 455
Db 424 CGAAGTCACCAACTAGCGGAGATCGTAGTAGAGCCGCCCTCTTTTATCTACCTACAC 483
Qy 456 aacgagcgaattaac 470
Db 484 AACGAGCGAATTAAAC 498

RESULT 15
AV749428 399 bp mRNA EST 19-OCT-2000
LOCUS AV749428 NPC Homo sapiens cDNA clone NPCBD10 5', mRNA sequence.
ACCESSION AV749428
VERSION AV749428.1 GI:10907276
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 399)
AUTHORS Song,H., Peng,Y., Gu,Y., Yang,Y., Gao,G., Xiao,H., Xu,X., Li,N.,
Qian,B., Liu,F., Du,J., Gao,X., Cheng,Z., Xu,Z., Zeng,L., Xu,S., Gu
,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Ye,M., Zhang,Q.,
Han,Z., Chen,Z., Hu,R. and Chen,J.
Homo sapiens NPC library cDNA clones
Unpublished (2000)
Contact: Qinghua Zhang
Shanghai Institute of Endocrinology, Rui-Jin Hospital
197 Rui-Jin II Road, Shanghai 200025, P. R. China
Tel: 86-21-64370045 (ex.663332)
Fax: 86-21-64743206
Email: mbsliens.sh.cn
This clone is available at Shanghai Hematology Institute in
Shanghai.
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang HI- Tech Park, Pudong.
LOCATION/Qualifiers
FEATURES
SOURCE 1..399
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="NPCBD10"
/clone_id="NPC"
/tissue_type="pituitary"
/dev_stage="Adult"
/lab_host="SOLR"
/note="Vector: Bluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT 75 a 113 c 92 g 113 t 6 others
ORIGIN
Query Match 38.5%; Score 360; DB 10; Length 399;
Best Local Similarity 96.3%; Pred. No. 1.8e-99;
Matches 388; Conservative 0; Mismatches 10; Indels 5; Gaps 2;

Qy 526 aatatgttcacatccggtgcacatcgtgtgtacggtgtgacgcgcatcctcaaaatgtgtc 585
Db 1 AATATGTTGCATCCGTCGTAATCATGCTGCTAACGTCGCAATTC---TGTGTC 56
Qy 586 tgcgattgacctctctgtgtgcatgacacacacacacacacacacacacacacacacacacac 645
Db 57 TGCATTCGATCTCTTGTGTGTCATGCACACACGAGCGCTCATACAGCGGCTTAACAGT 116
Qy 646 gcgtgaccagtggtgtggtgaagttgtggtatgacatcgtcacagcgatagctgtc 705

```

```

Db 117 GCCTGACCAAGGTGGTGGTGAAGGTTTGGAATTACATCTGCACAGCCGATATGCTGC 176
Qy 706 gcttgctgacatccttgaaatagccgacgaccttgcaatcttcggaactcttcgacaac 765
Db 177 GCTTGCTGGCATCTTGAATAGCCGACGCTTTGCAATCTTCGCACTCTTTTCGACAC 236
Qy 766 tctccccaacagctcgttttgcaatataaac-cgacagcctgtaccatggcaatctc 824
Db 237 TCTCCCAACAGCTCTGTTTGCAATATMAACNCGACGGCTGTACCATGSMATCTC 296
Qy 825 tgcatttgcctcccgagcgtgcgcgaactcagcaataatccgataagcgaatgttcga 884
Db 297 TGCATCTTGCCCGGCGCTGCGGCACTAGCGCAATMAATCCCATMAAGCMAATGTTGMA 356
Qy 885 gcaattgacagcttgcttgcttgatgattatcttcctcaagctgccc 927
Db 357 GCATTCGATGACTTGTGCTTAGTATTTCTTCCTTCGCTTGGCCC 399

```

Search completed: January 8, 2002, 15:13:30  
Job time: 4806 sec



## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1000	100.0	11933	6	AR138377	AR138377 Sequence
2	1000	100.0	11933	6	AR146887	AR146887 Sequence
3	725.2	72.5	4999	6	AR138378	AR138378 Sequence
4	725.2	72.5	4999	6	AR146888	AR146888 Sequence
5	577.4	57.7	5035	6	AR014084	AR014084 Sequence
6	577.4	57.7	5035	6	AR071306	AR071306 Sequence
7	577.4	57.7	5035	6	AR112722	AR112722 Sequence
8	575	57.5	8241	6	A05328	A05328 Synthetic F
9	575	57.5	8241	6	A07042	A07042 Artificial
10	575	57.5	8831	6	E00527	E00527 Complete CD
11	575	57.5	8967	6	AX052730	AX052730 Sequence
12	575	57.5	8967	6	AX052730	AX052730 Sequence
13	575	57.5	8967	6	HSFV11IR	HSFV11IR
14	575	57.5	9029	6	HSFV11IR	HSFV11IR
15	574	57.4	7272	6	HOVEV11IC	HOVEV11IC
16	573.4	57.3	9009	6	I05404	I05404
17	573.4	57.3	9009	6	AR003710	AR003710 Sequence
18	573.4	57.3	9009	6	AR029065	AR029065 Sequence
19	573.4	57.3	9009	6	AR126884	AR126884 Sequence
20	573.4	57.3	9009	6	I31901	I31901 Sequence
21	573.4	57.3	9009	6	163424	163424 Sequence
22	572.4	57.2	9056	6	K01740	K01740 Human coagu
23	570.2	57.0	7440	6	HOVEV11II	HOVEV11II
24	567.8	56.8	9354	6	I27063	I27063 Sequence
25	563.2	56.2	7440	6	I08345	I08345 Sequence
26	517	51.7	4278	6	AR003585	AR003585 Sequence
27	517	51.7	4281	6	I08457	I08457 Sequence
28	517	51.7	4281	6	I08643	I08643 Sequence
29	517	51.7	4281	6	I08642	I08642 Sequence
30	517	51.7	4670	6	I08641	I08641 Sequence
31	517	51.7	6599	6	AR110040	AR110040 Sequence
32	501.4	50.1	1933	6	I77105	I77105 Sequence
33	444.2	44.4	7032	6	I02047	I02047 Sequence
34	443	44.3	7493	6	AF016234	AF016234 Canis fam
35	443	44.3	7493	6	AR003712	AR003712 Sequence
36	443	44.3	7493	6	AR029067	AR029067 Sequence
37	443	44.3	7493	6	AR012686	AR012686 Sequence
38	443	44.3	7493	6	I63427	I63427 Sequence
39	442.6	44.3	7145	4	M05573	M05573 Mus domest
40	414.8	41.5	6539	4	AF049489	AF049489 Canis fam
41	413.8	41.4	4334	6	SS049517	SS049517 Sus scrofa
42	413.8	41.4	4334	6	AR029098	AR029098 Sequence
43	413.8	41.4	4334	6	AR126917	AR126917 Sequence
44	413.8	41.4	6402	6	AR029097	AR029097 Sequence
45	371.8	37.2	553	4	AF180523	AF180523 Ovis arie

[illegible]



```

|||||
Db 661 AACATGGCTAAGCAAGGCCACCTGAGMGGTCTGTAGAGTCTTACCATCCAGCTGAG 720
Qy 721 gttatgatagcagtgatcacttaagaacatgggtccatccctgtaagcttcat 780
Db 721 GTTTATGATACAGTGTGCTACTTACCTTAAGAACAGGCTCTCCATCCTGTAGCTTTCAT 780
Qy 781 gctgttgatcctcactcaggaaagcttcgagagctgaatgatgatcaagcaagt 840
Db 781 GCTGTGTTGTTATCTTACTGCAAGAGCTTGTGAGGAGCGATATGATGATGACAGCAGT 840
Qy 841 caaaggagaaagaagaatgaatgaatcttcctcctggtggaagcacaataatgctggag 900
Db 841 CAAGAGGAGAAAGAAATGATGAAGAGCTTCTCCGTGTGAGGAGCCATATATGTGTGGCAG 900
Qy 901 gtcctgaagaagaatggtcaaatggcttcgagccactgctgcttaactacatcatct 960
Db 901 GTCCTGAAGAAGATGATGCTCAATGGCTCTGACCCACTGTGCTTACTACTATATCTT 960
Qy 961 tctcatgtgacctggtlaaagaacttgaaatltcaagcctca 1000
Db 961 TCTCATGTGACCTGCTGAAGAGCTTGAATTCAGGCTCA 1000

```

```

RESULT 3
ARI38378 ARI38378 4999 bp DNA PAT 16-JUN-2001
LOCUS Sequence 14 from patent US 6200560.
ACCESSION ARI38378
VERSION ARI38378.1 GI:14480723
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 4999)
AUTHORS Couto, L.B., Colosi, P.C. and Qian, X.
TITLE Adeno-associated virus vectors for expression of factor VIII by
JOURNAL Patent: US 6200560-A 14 13-MAR-2001;
FEATURES Location/Qualifiers
source 1..4999
BASE COUNT 1375 a 1151 c 1155 g 1318 t
ORIGIN

```

```

Query Match 72.5%; Score 725.2; DB 6; Length 4999;
Best Local Similarity 85.9%; Pred. No. 6,7e-210;
Matches 859; Conservative 0; Mismatches 118; Indels 23; Gaps 4;

```

```

Qy 1 cagctgcgcgtcgtcgtcactgagagccgcgggaaagcccgagcgttcgggagacc 60
Db 13 CAGCTGCGCGCTCGCTCCTCCTCAGTAGGCGCCGCGGAAAGCCCGGCGTGGCGCACC 72
Qy 61 ttgtgtcgcgcgcgtcgtcagtgagcagcagcagcagagagagagtggtccaaactcatc 120
Db 73 TTTGTGTCGCCCGCTCAGTATAGCAGGAGGCGCAGAGAGGAGTGCGCAACTCCATC 132
Qy 121 actaagggttcctcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 180
Db 121 ACTAGGCGGTTCTCGCGCGCAGCGCTGTGTGCGCGGCGGTAAA-----CTGGGAAGT 185
Qy 181 ttgttctaataacatcagaggaatgttgttcttaataacatcactacagttatglt 240
Db 181 GATGTGCTGATGCTGCTCC---GCCTTTTCGCGAGGGTGCGGCGAGAACCGTATATAGT 242
Qy 241 aaagaatlatatagcagagcttcttcgcacacagatcaacttcgcggggtgcgcgcct 300
Db 243 GCAGTATGTCGCGTGAACGTTCTTTTCGCAACGGGTT-----TGGCGCCCG 290
Qy 301 aggcaggttaagtcggtggtgtgttccgcgggcctgctcttlaaggtatagccct 360
Db 291 CGGCGAGTATGTCGCAAGGAGTGTGTTCTTAATACATGCTGCTCAAGGAGATGTGTTGT 350

```

```

Qy 361 tgcgtgccttgtaattactgacactgacatccacttttcttctcccaagatcat 420
Db 351 TCTTAATATCCATATGATGACACTGACATCCACTTTTCTTTTCTCCACAGATATGCA- 409
Qy 421 tccacctgcaaatagagctccaccctgctcttcttcgtccctttgagcttgcctt 480
Db 410 TCCACATGCAAAATAGAGCTCTCCACCTGCTCTTCTGTGCTTTTGGATGCTGCTTT 469
Qy 481 agtgcacacagaagataactaactggtgtgagtggaactgtaatggaactatagaaat 540
Db 470 AGTGCACACAGAAAGATATACCTGCTGATGAGTACTGTATGAGACTATATGCAAAAT 529
Qy 541 gatctgggtgagctgctgctgagcgaagatctcctcctcagatgagtgccaaatcttcca 600
Db 530 GATCTGCTGAGCTGCTGCTGAGCAGCAATTTCTCCCTAGAGTGCCAAAAATCTTTTCCA 589
Qy 601 ttaacaacctcagtcgtglaaaaaaagactctgttctgtagaatltcaagatcaccttlt 660
Db 590 TTCACACCTCACTCGTGTACAAAAGACTGTGTGTGAAATTCACGATACCTTTTC 649
Qy 661 aacatgctaaagcgaagccacccttgatgggtctgctagtcctaccatccagcgtgag 720
Db 650 AACATGCTAAGCCAAAGGCCACCTGATGGTCTGTAGTCTATGCTACATCCAGGCTGAG 709
Qy 721 gttatgatacagtggtcatlaacttaagaacatggttcccatctgctcaagttcat 780
Db 710 GTTTATATATACAGTGCATATACCTTAAGAACATGCTTCCATCTGTGCTTCTCAT 769
Qy 781 gctgtgtgtatcctcctgagaaagcttcctgagagcgctgaatagatagatagaccagt 840
Db 770 GCTGTGCTGTATCCTACTGTGAAGCTTCTGAGGAGCTGATATATATATATGACAGCACT 829
Qy 841 caaaggagaaagaagaatgataagcttccctggtgtggaagccatacatatgctggag 900
Db 830 CAAGGAGAAAGAAATGATGAATGAATCTTCCCTGAGGAGCCATATATGCTTGGCAG 889
Qy 901 gtcctgaagaagaatggttccaaatgctcctgagccactgctgcttaactacatcatct 960
Db 890 GTCCGAAAGAGATGCTGCAATGCTCTGACCACTGTGCTTACTTACTATATCTT 949
Qy 961 tctcatgtgacctggtlaaagaacttgaaatltcaagcctca 1000
Db 950 TCTCATGTGACCTGCTGAAGAGCTTGAATTCAGGCTCA 969

```

```

RESULT 4
ARI46888 ARI46888 4999 bp DNA PAT 08-AUG-2001
LOCUS Sequence 14 from patent US 6221349.
ACCESSION ARI46888
VERSION ARI46888.1 GI:15110691
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 4999)
AUTHORS Couto, L.B., Colosi, P.C. and Qian, X.
TITLE Adeno-associated vectors for expression of factor VIII by target
JOURNAL Patent: US 6221349-A 14 24-APR-2001;
FEATURES Location/Qualifiers
source 1..4999
BASE COUNT 1375 a 1151 c 1155 g 1318 t
ORIGIN

```

```

Query Match 72.5%; Score 725.2; DB 6; Length 4999;
Best Local Similarity 85.9%; Pred. No. 6,7e-210;
Matches 859; Conservative 0; Mismatches 118; Indels 23; Gaps 4;
Qy 1 cagctgcgcgtcgtcgtcactgagagccgcgcgggcaagcccggtcgtcgggagacc 60

```









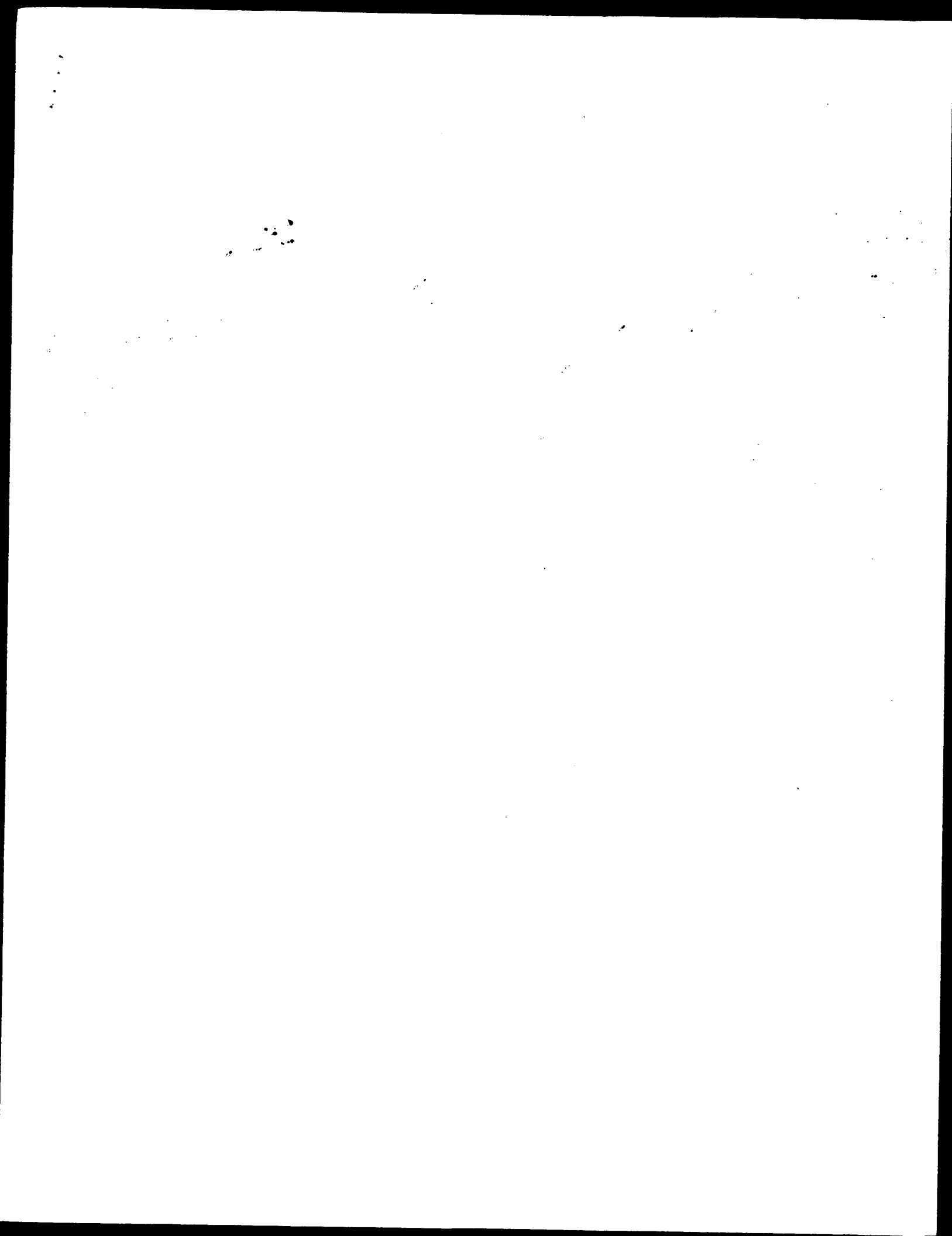








	/note="C an wt: t in hemophiliac [3]"
BASE COUNT	2860 a 1910 c 1848 g 2411 t
ORIGIN	185 bp upstream of SacI site.



(without alignment), 1999.219 Million cell updates/sec

```

Title:          US-09-740-211_13_COPY1_1_1000
Perfect score: 1000
Sequence:       1 cagctgcgcgcctgcctgcct.....agacttgatcaggcctca 1000
Scoring table: IDENTITY_NUC
                  Gapop 10.0 , Gapext 1.0

```

Searched: 930621 segs, 428662619 residues  
Total number of hits satisfying chosen parameters: 1861243

Minimum DB seq length:	0
Maximum DB seq length:	2000000000

```

post-processing: Minimum Match, 0%
                  Maximum Match, 100%
                  Listing first 45 summaries

```

```

Database :
1: N_Geneseq_1101.*
2: /SIDS2/gcgdata/geneseq/geneseqn/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseqn/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseqn/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseqn/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseqn/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseqn/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseqn/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseqn/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseqn/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseqn/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseqn/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseqn/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseqn/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseqn/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseqn/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseqn/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT.*

```

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1000	100.0	11933	21	AAPO00121	Recombinant adeno
2	1000	100.0	11933	22	AAD008612	Human factor VIII
3	725.2	72.5	4999	21	AAPO00122	Recombinant adeno
4	725.2	72.5	4999	22	AAAD08613	Human factor VIII
5	617.4	61.7	7944	22	AAAP84647	Plasmid DLZ6 enco
6	577.4	57.7	.5035	18	AAAT98811	Factor VIII -db695
7	577.4	57.7	12445	21	AAAS92259	Factor VIII H5GRetw
8	575.8	57.6	9164	20	AAAS82232	Beta-domain delet
9	575.8	57.6	11846	20	AAAS82261	Factor VIII prote
10	575.8	57.6	12022	20	AAAS82260	Factor VIII prote
11	575	57.5	4832	19	AAV19581	Human factor VIII

12	575	57.5	4832	19	AAV15338	Human Factor VIII
13	575	57.5	6300	17	AAT03571	Factor VIII cDNA.
14	575	57.5	8241	9	AAH81439	Factor VIII cDNA. i
15	575	57.5	8241	9	AAH81096	cDNA sequence enco
16	575	57.5	8967	17	AAT31031	Factor-VIII full-l
17	575	57.5	8967	22	AACB7526	Human factor VIII
18	575	57.5	8975	6	AAAN0054	Human factor VIII
19	575	57.5	8975	21	AAAG38604	Human full-length
20	575	57.5	9029	22	AAAF60309	Human factor VIII
21	575	57.5	9068	19	AAV153359	Human factor VIII
22	575	57.5	9080	19	AAV19580	Human factor III e
23	574	57.4	4830	9	AAH81544	Human Factor VIII-
24	574	57.4	7053	18	AAE51357	Factor VIII::C codl
25	573.4	57.3	9009	14	AAAG01285	Human Factor VIII
26	573.4	57.3	9009	18	AAE61548	Factor VIII::C (Arg
27	573.4	57.3	9009	19	AAAV25810	Human factor VIII
28	573.4	57.3	9009	20	AAAG11162	Human factor VIII
29	573.4	57.3	9009	22	AAAC95058	Human factor VIII
30	572.4	57.2	5094	21	AAAG49231	Human factor VIII
31	572.4	57.2	7056	15	AAAG66653	DNA construct HSO/
32	570.8	57.1	4629	19	AAH88293	Sequence of human
33	570.2	57.0	9009	20	AAV18884	Human Factor VIII
34	569.2	56.9	4629	16	AAAG6016	Human Factor VIII
35	568.6	56.9	7440	7	AAAG0689	Homo sapiens Factor
36	567.8	56.5	9354	18	AAT73164	B-domain deleted F
37	564.8	55.8	4373	20	AAH82258	Sequence encoding
38	552.6	55.3	7440	6	AAAS0375	cDNA encoding huma
39	517	51.7	4273	9	AAH80445	Beta-domain delete
40	517	51.7	4273	9	AAH80445	Modified factor VI
41	517	51.7	4275	10	AAH80445	Modified factor VI
42	517	51.7	4545	9	AAH80444	DNA encoding 740 A
43	513.8	51.4	4670	19	AAAB23339	Modified factor VI
44	506.2	50.6	9009	19	AAV12112	Human Factor-VIII
45	486.2	48.6	7914	22	AAAF84548	Homo sapiens Factor
						Plasmid DL7 encod

RESULT	1	
AD000121		
ID	AD000121	standard; DNA; 11933 BP.
XX		
AC	AD000121;	
DT		
DT	31-JUL-2000	(first entry)
DE		
XX	Recombinant adeno associated vector construct, pAAV-F8-1.	
KW	Recombinant Adeno Associated Vector; rAAV; pAAV-F8-1; human Factor VIII;	
KW	hvtIII; hNF-3 albumin promoter; human elongation factor-1alpha; EFlalpha;	
KW	human growth hormone; hGH; inverted terminal repeat; ITR; haemophilia;	
KW	gene therapy; ds.	
OS	Adeno associated virus.	
XX		
PN	WC200023116-A1.	
XX		
PD	27-APR-2000.	
XX		
PF	19-OCT-1999;	99WO-US24495.
XX		
XX	20-OCT-1998;	98US-0104994.
PR	24-MAR-1999;	98US-0125974.
PR	30-JUL-1999;	99US-0364862.
XX		
PA	(AVIG-) AVIGEN INC.	
XX		
PI	Couto LB, Colosi PC;	
XX		
DR	WPI; 2000-339536/29.	
XX		
PT	New recombinant adenovirus-associated vector, useful for gene therapy	

PT to treat hemophilia, comprises at least a portion of Factor VIII  
 operably linked to control sequence -  
 Example 2, Fig 5; 92pp; English.

The present DNA sequence is a recombinant adenovirus-associated vector, (AAV) construct, pAAV-F8-1. This expression vector comprises the HNF-3 albumin promoter, the first intron (-573 to +985) of human elongation factor-1alpha (EF1alpha) gene, human Factor VIII coding sequence (hFVIII) and polyadenylation signal from human growth hormone (hGH). This sequence is inserted between the AAV inverted terminal repeat (ITR) regions. The hFVIII coding region comprises the heavy chain gene segment with the first 57 bp encoding the 19 amino acid signal peptide, A1 and A2 domains and 5 amino acids from the N-terminus of the B domain. The light chain segment comprises the C-terminal 85 amino acids of B domain and the A3, C1 and C2 domains. Both the heavy and light chain segments are cloned into the same plasmid separated by 42 nucleotides coding for 14 residues of the B domain, that is deleted. This plasmid is operably linked to Factor VIII gene. The adeno-associated viral vectors are used for gene therapy to treat haemophilia. This method allows prolonged expression of CC therapy, because of their broad host range, safety profile and duration of expression in the infected hosts.

Sequence 11933 BP; 328 A; 2818 C; 2717 G; 3140 T; 0 other;

Query Match 100.0%; Score 1000; DB 21; Length 11933;  
 Best Local Similarity 100.0%; Pred. No. 9.8e-278;  
 Matches 1000; Conservative: 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 cagctggcgcgtgcgtgcctacgtgagccgcccgaagcccgccgctgcgcgcacc 60
DB 1 cagctggcgcgtgcgtgcctacgtgagccgcccgaagcccgccgctgcgcgcacc 60
QY 61 ttgtgtcccccgcgcctcctgagcagcagcagcagcagcagcagcagcagcagc 120
DB 61 ttgtgtcccccgcgcctcctgagcagcagcagcagcagcagcagcagcagcagc 120
QY 121 actaggaggttcctgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 120
DB 121 actaggaggttcctgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 120
QY 122 actaggaggttcctgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 180
DB 122 actaggaggttcctgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 180
QY 181 ttgttcttaataacatccagcagcagcagcagcagcagcagcagcagcagcagc 240
DB 181 ttgttcttaataacatccagcagcagcagcagcagcagcagcagcagcagcagc 240
QY 241 aagcagatatttgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 300
DB 241 aagcagatatttgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 300
QY 301 aggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 360
DB 301 aggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 360
QY 361 tgcgtgccttgatctgacatcagcagcagcagcagcagcagcagcagcagcagc 420
DB 361 tgcgtgccttgatctgacatcagcagcagcagcagcagcagcagcagcagcagc 420
QY 421 tccacatgcaaatagagctccacacgtctcttccttccttccttccttccttc 480
DB 421 tccacatgcaaatagagctccacacgtctcttccttccttccttccttccttc 480
QY 481 agtgcacacagagatatactactggtgctgagcagcagcagcagcagcagcagcag 540
DB 481 agtgcacacagagatatactactggtgctgagcagcagcagcagcagcagcagcag 540
QY 541 gatctcgtgtgagctgctgctgagcagcagcagcagcagcagcagcagcagcagc 600
DB 541 gatctcgtgtgagctgctgctgagcagcagcagcagcagcagcagcagcagcagc 600
QY 601 ttcaacacctcagctcgtgtgtcaaaagactgtttgtagaatcaacgagcagcagc 660

```

```

DB 601 ttcaacacctcagctcgtgtgtcaaaagactgtttgtagaatcaacgagcagcagc 660
QY 661 aacatcgtcgaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 720
DB 661 aacatcgtcgaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 720
QY 721 gttatgatacagtgatcattacactaaagaacatggtctccatcccgtaagctcat 780
DB 721 gttatgatacagtgatcattacactaaagaacatggtctccatcccgtaagctcat 780
QY 781 gctgtgtgtatccctcagcagcagcagcagcagcagcagcagcagcagcagcagc 840
DB 781 gctgtgtgtatccctcagcagcagcagcagcagcagcagcagcagcagcagcagc 840
QY 841 caaaggagagaaagatataagctcttcctcctggtgagcagcagcagcagcagcag 900
DB 841 caaaggagagaaagatataagctcttcctcctggtgagcagcagcagcagcagcag 900
QY 901 gtcctgaagagagatggtccatgagcagcagcagcagcagcagcagcagcagcagc 960
DB 901 gtcctgaagagagatggtccatgagcagcagcagcagcagcagcagcagcagcagc 960
QY 961 tctcatgtgacctggtgaaagacttgatcagcagcagcagcagcagcagcagcagc 1000
DB 961 tctcatgtgacctggtgaaagacttgatcagcagcagcagcagcagcagcagcagc 1000

```

RESULT 2  
 AAD08612 standard; DNA; 11933 BP.

AC AAD08612;

DT 04-SEP-2001 (first entry)

Human factor VIII expressing rAAV vector pAAV-F8-1 partial sequence.  
 Recombinant adeno-associated viral vector; rAAV; factor VIII; FVIII;  
 blood clotting disorder; gene therapy; haemophilia A; human;  
 pAAV-F8-1; ds.

OS Chimeric - Adeno associated virus.

OS Chimeric - Mus sp.

OS Chimeric - Homo sapiens.

OS Chimeric - Oryctolagus cuniculus.

PN W020014510-A1.

PD 28-JUN-2001.

PF 21-DEC-2000; 2000WO-US34925.

PR 22-DEC-1999; 99US-0470618.

PA (AVIGEN) AVIGEN INC.

PI Couto LB, Colosi PC, Qian X;

PI WPI: 2001-417955/44.

DR Treating blood clotting disorder, especially hemophilia in mammals, by

administering recombinant adeno-associated vectors which express blood

coagulation factor VIII

PS Claim 18; Fig 5; 90pp; English.

PT The present invention relates to a method for treating a subject

suffering from a blood clotting disorder. The method comprises

administering a recombinant adeno-associated virion (rAAV) comprising

a nucleotide sequence encoding the light chain of factor VIII and a

second recombinant adeno-associated virion comprising a nucleotide

sequence encoding the heavy chain of factor VIII. The rAAV vector is







Db	133	actagagggttccttcgagcgacacgctggtgagcgaggtaaa-----ctggaaaaat	185
Qy	181	ttgtcttaataacatctccagggaatggttcttctaataatacatcatcagttatgtt	240
Db	186	gagtcgtgctactgctcc-----gccttttccggagggtggtgggagacgatatataagt	242
Qy	241	aaagaagattattagagcgagctcttcttgacaacagatcatcacttccggtgccccct	300
Db	243	gcagtagtcgcgcgtgacgctcttcttcgcaacgggtt-----ggccggccccg	290
Qy	301	agcgaggaatgagtcgcgtgtgtgtgttcccgaggcgctgcctcttaaggttatggccct	360
Db	291	cggaggaatgagtgccgagggaatggttcttctaataatacactcctccagggaatgttgt	350
Qy	361	tgcgtgctggaattactctaacacttaacatccactttcttcttccacagtgatcat	420
Db	351	tcttaataatacatctactgacactgacatcatcactttcttcttccacaagatcatga-	409
Qy	421	tccacactgacaataagacgtctccaccgtcttcttcttgagctcttcttgagatctgtctt	480
Db	410	tccacactgacaataagacgtctccaccgtcttcttcttgagctcttcttgagatctgtctt	469
Qy	481	agtgccacccaagaatactactacccctggtgtgagatggaactgtcatatggagactatgacaagt	540
Db	470	agtgccacccaagaatactactacccctggtgtgagatggaactgtcatatggagactatgacaagt	529
Qy	541	gacttcgtgtgagctgctgtgtgagcaagaattcctctctagaatgagcacaatttttcca	600
Db	530	gacttcgtgtgagctgctgtgtgagcaagaattcctctctagaatgagcacaatttttcca	589
Qy	601	ttcaacacccctcagtcggtgacaanaaagaactctgtttgtagaatcaagatcacacttttc	660
Db	590	ttcaacacccctcagtcggtgacaanaaagaactctgtttgtagaatcaagatcacacttttc	649
Qy	661	aacatcgtctaagccaagagccacccctggaatgggtctgtctaggttcttaccatccagctgag	720
Db	650	aacatcgtctaagccaagagccacccctggaatgggtctgtctaggttcttaccatccagctgag	709
Qy	721	gtttatgatacagtgatcatcttaccattagaacaatggtcttccatccctgtcagttctcat	780
Db	710	gtttatgatacagtgatcatcttaccattagaacaatggtcttccatccctgtcagttctcat	769
Qy	781	gctgtgtgagatctactacttggaagctctctggaggagcttgataatgatactgacagcagt	840
Db	770	gctgtgtgagatctactacttggaagctctctggaggagcttgataatgatactgacagcagt	829
Qy	841	caaaaggagaagaagatgataaaggtcttccctcgtgtgaaagccatacatatgtctggcag	900
Db	830	caaaaggagaagaagatgataaaggtcttccctcgtgtgaaagccatacatatgtctggcag	889
Qy	901	gtctcgaagaagaatggttccaatgtgctctggaccacacttgcttactactatcatct	960
Db	890	gtctcgaagaagaatggttccaatgtgctctggaccacacttgcttactactatcatct	949
Qy	961	tctcatgtgacctggttaaaagacttgaaatctcaggctctca	1000
Db	950	tctcatgtgacctggttaaaagacttgaaatctcaggctctca	989
RESULT 5			
AAFB4647			
ID AAFB4647 standard; DNA: 7944 BP.			
XX			
AAF84647;			
XX			
DT 29-JUN-2001 (first entry)			
XX			
DE plasmid DL26 encoding human B-domain deleted factor VIII.			
XX			
KW Adeno-associated virus vector; B-domain; factor VIII; haemophilia A;			
KW coagulation disorder; ss. ;			
XX			

OS	Synthetic.
OS	Homo sapiens.
OS	Hepatitis B virus.
XX	
FH	Key
FT	misc_feature
FT	Location/Qualifiers
FT	1..146
FT	/tag= "a"
FT	/note= "inverted terminal repeat"
FT	150..278
FT	/tag= "b"
FT	/note= "hepatitis B virus EnhI enhancer"
FT	420..4835
FT	/tag= "c"
FT	/note= "human B-domain deleted factor VIII"
FT	4840..4914
FT	/tag= "d"
FT	/note= "HBV polyA sequence"
FT	4916..5084
FT	/tag= "e"
FT	/note= "Inverted terminal repeat"
FT	
FN	misc_feature
PD	
PD	W0200127303-A1.
XX	
PN	
PD	19-APR-2001.
XX	
PE	12-OCT-2000; 2000MO-US28221.
XX	
PR	12-OCT-1999; 99US-0158780.
XX	
XX	(UNNC-) UNIV NORTH CAROLINA.
PA	
XX	
XX	Walsh CE, Chao H, Burstein H, Lynch CM, Stepan AM, Munson K;
XX	
DR	WPI: 2001-273781/28.
XX	P-SDB; AAB67959.
PT	New recombinant adeno-associated virus vector useful for treating
PT	haemophilia A, comprises heterologous nucleotide sequence encoding
PT	B-domain deleted human factor VIII operably linked with liver-preferred
PT	expression control element -
XX	
PS	Claim 64; Fig 1; 87pp; English.
XX	
CC	The specification describes a recombinant adeno-associated virus (rAAV)
CC	vector. The vector comprises a heterologous nucleotide sequence
CC	encoding B-domain deleted factor VIII operably linked with at least one
CC	enhancer and at least one promoter. The method results in the production
CC	of high titer rAAV vector stocks carrying the B-domain deleted factor
CC	VIII transgenes and expression cassettes, which generate adequate titers
CC	of virus for in vivo administration. The recombinant vectors are useful
CC	for treating haemophilia A, where the liver expresses the encoded
CC	B-domain deleted factor VIII, which is secreted into the blood. They are
CC	also useful for the treatment of other coagulation disorders. The
CC	present sequence encodes a B-domain deleted factor VIII.
XX	
SO	Sequence 7944 BP; 2142 A; 1902 C; 1909 G; 1991 T; 0 other;
OY	
D6	Query Match 61.7%; Score 617.4; DB 22; Length 7944;
OY	Best Local Similarity 78.5%; Pred. No. 1.4e-167; Indels 18; Gaps 3;
OY	Matches 783; Conservative 0; Mismatches 196;
OY	
OY	4 ctgcgctcgctcgctcaatgaagccggcgcaaagcccggtcggtcgacttt 63
OY	
OY	15 ctgcgcgtctgcgtcgtacgaaggccggcggaagaagtgcgcaattcatcact 74
OY	
OY	64 gtttcgccgcctcagtgagcgagcgagcgcgagaagaagatgtgccaattcact 123
OY	
OY	75 gcccgagcgagcctcagtgagcgagcgagcgaggaagaaatgtgccaattcact 134
OY	
OY	124 agggatcttcgcggcgcccgcaaggaaatgttgttctaataaccatccagggaatcttg 183
OY	
OY	135 agggatcttc-----cagatctctttaagtaaacagtatcagaaccttt 180
DB	

XX	
PD	22-MAY-1997.
XX	
PF	13-NOV-1996; 96WO-EP04977.
XX	
PR	13-NOV-1995; 95US-0558107.
XX	
PA	(IMMO ) IMMUNO AG.
XX	
P1	Voorberg JJ;
XX	
DR	WPI: 1997-289291/26.
DR	P-PSDB; AAM18670.
XX	
XX	
PT	Hybrid Factor VIII with modified activity, comprises region from
PT	donor anticoagulant or antithrombotic protein - useful for treatment
XX	
XX	
S	Claim 16; Page 52-60; 96pp; English.

[illegible]

		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	
--	--	---	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	--

OY 782 ctgttggtatctactgtgaagctctgaaggagctgaatatgtatgcagaccagtc 841  
 CC ctgttggtatctactgtgaagctctgaaggagctgaatatgtatgcagaccagtc 841  
 DB 390 ctgttggtatctactgtgaagctctgaaggagctgaatatgtatgcagaccagtc 449  
 OY 842 aaaggagaaagagatgataaaatctctccctggtggaagccataatgctgcagc 901  
 CC aaaggagaaagagatgataaaatctctccctggtggaagccataatgctgcagc 901  
 DB 450 aaaggagaaagagatgataaaatctctccctggtggaagccataatgctgcagc 509  
 OY 902 tccgtgaagaagaatgctccatctgctcctcagccacacgtgcttactactatctt 961  
 CC tccgtgaagaagaatgctccatctgctcctcagccacacgtgcttactactatctt 961  
 DB 510 tccgtgaagaagaatgctccatctgctcctcagccacacgtgcttactactatctt 569  
 OY 962 ctcatgtgacctggttaaaagacttgaattcagccctca 1000  
 CC ctcatgtgacctggttaaaagacttgaattcagccctca 1000  
 DB 570 ctcatgtgacctggttaaaagacttgaattcagccctca 608  
 RESULT 7  
 ID AAA49232 standard; DNA; 12445 BP.  
 AC AAA49232;  
 DT 26-SRP-2000 (first entry)  
 DE Vector HSQReNeo for transforming endothelial cells.  
 XX Hemostatic; antianemic; antidiabetic; neurotropic; neuroprotective; vector;  
 KM osteopathic; antisticking; immunostimulant; gene therapy; collagen;  
 KM endothelial cell; peripheral blood; buffy coat cell; VEGF; PCR primer;  
 KM vascular endothelial growth factor; bovine brain extract; hemophilia;  
 KM factor VIII; human; transgene; adenosine deaminase deficiency; ss;  
 KM sickle cell anemia; thalassemia; diabetes; alpha-antitrypsin deficiency;  
 KM Alzheimer's disease; brain disease; heart disease; immune system defect;  
 KM bone fracture; osteoporosis.  
 XX Homo sapiens.  
 OS Synthetic.  
 PN W0200032750-A1.  
 PD 08-JUN-2000.  
 PP 24-NOV-1999; 99MO-US28033.  
 PR 24-NOV-1998; 98US-0109687.  
 PA (MNU) UNIV MINNESOTA.  
 PA (UEM-) UNIV EMORY.  
 PA (HEB-) HEBBEL R P.  
 PA (LINV) LIN Y.  
 PA (LOLL) LOLLAR J S.  
 PI Heibel RP, Lin Y, Lollar JS.  
 DR WPI: 2000-412303/35.  
 XX Expanding population of endothelial cells useful to biocompatibilize  
 PT implantable medical devices comprises contacting buffy coat cells with  
 PT collagen I coated surface in culture medium comprising vascular  
 PT endothelial growth factor  
 PS Claim 19; Fig 4; 53pp; English.  
 XX The invention relates to a method for expanding the population of  
 CC endothelial cells (EC) obtained from peripheral blood, by culturing,  
 CC in contact with a collagen I coated surface, buffy coat cells obtained  
 CC from peripheral mammalian blood in the presence of a culture medium  
 CC containing vascular endothelial growth factor (VEGF) and free of bovine  
 CC brain extract. EC are useful for treating hemophilia by introducing it  
 CC into the blood stream of a mammal, so that an effective amount of  
 CC factor VIII protein is secreted in the blood stream of the mammal.

CC This sequence represents the vector HSQReNeo where the human factor  
 CC VIII gene into which an enhanced green fluorescent protein coding  
 CC sequence (H8Q) has been inserted, is subcloned. Transgenic EC transduced  
 CC in vitro are useful for improving prosthetic implants. EC is also useful  
 CC for diagnosing clotting disorders where indication of disease is  
 CC associated with a reduction in the activity of an enzyme. EC is also  
 CC useful in gene therapy for treating the variety of diseases including  
 CC adenosine deaminase deficiency, sickle cell anemia, thalassemia,  
 CC hemophilia, diabetes, alpha-antitrypsin deficiency, brain disorders  
 CC such as Alzheimer's disease, heart diseases, defects in immune system,  
 CC for repairing bone fractures and to treat or prevent osteoporosis.  
 XX Sequence 12445 BP; 3255 A; 2989 C; 2951 G; 3250 T; 0 other;

Query Match 57.7%; Score 577.4; DB 21; Length 12445;  
 Best Local Similarity 99.8%; Pred. No. 5.8e-156;  
 Matches 578; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 422 ccacatgcaaatagagctccacactgtctctcttctgtgaccttgcagttcgtctta 481  
 DB 763 ccacatgcaaatagagctccacactgtctctcttctgtgaccttgcagttcgtctta 822  
 OY 482 gtcgacccagaagatactactcctggtgcagtggaacttcagtggaactatgcagagtg 541  
 DB 823 gtcgacccagaagatactactcctggtgcagtggaacttcagtggaactatgcagagtg 882  
 OY 542 atctcggtgagctgctggtgacgcaagattcctcctcagagtgccaaatcttccat 601  
 DB 883 atctcggtgagctgctggtgacgcaagattcctcctcagagtgccaaatcttccat 942  
 OY 602 tcaacacctcagtcgtgtacaaaagacctgtttgtagaattcagcagatcccttca 661  
 DB 943 tcaacacctcagtcgtgtacaaaagacctgtttgtagaattcagcagatcccttca 1002  
 OY 662 acatcgtaagcacaagccacccctgtgatgctcgtcgtacgtccacatccagctgag 721  
 DB 1003 acatcgtaagcacaagccacccctgtgatgctcgtcgtacgtccacatccagctgag 1062  
 OY 722 ttatgatacagtggtcattacacttaagaacctggtccatccctgtcagttcagtg 781  
 DB 1063 ttatgatacagtggtcattacacttaagaacctggtccatccctgtcagttcagtg 1122  
 OY 782 ctgttggtatctactgtgaagctctgaaggagctgaatatgtatgcagaccagtc 841  
 DB 1123 ctgttggtatctactgtgaagctctgaaggagctgaatatgtatgcagaccagtc 1182  
 OY 842 aaaggagaaagagatgataaaatctctccctggtggaagccataatgctgcagc 901  
 DB 1183 aaaggagaaagagatgataaaatctctccctggtggaagccataatgctgcagc 1242  
 OY 902 tccgtgaagaagaatgctccatctgctcctcagccacacgtgcttactactatctt 961  
 DB 1243 tccgtgaagaagaatgctccatctgctcctcagccacacgtgcttactactatctt 1302  
 OY 962 ctcatgtgacctggttaaaagacttgaattcagccctca 1000  
 DB 1303 ctcatgtgacctggttaaaagacttgaattcagccctca 1341  
 RESULT 8  
 ID AAX82259 standard; CDNA; 9164 BP.  
 AC AAX82259;  
 DT 18-AUG-1999 (first entry)  
 DE Beta-domain deleted Factor VIII protein encoding gene (uncorrected).  
 XX Factor VIII protein; gene modification; gene therapy; clinical disorder;  
 KM splicing pattern; RNA processing; gene regulation; beta-domain; human;  
 KM ss.





QY 782 cgttggtgtagtaccctggaagctcttgagagagcgaatgatgacgaagcgc 841  
 CC ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 CC the RRV vector preparation. RRV's can be used for in vivo delivery of  
 CC therapeutic protein to treat, e.g. haemophilia A, haemophilia B,  
 CC thrombosis, hypercoagulable disorders, liver diseases such as hepatitis,  
 CC disorders such as thalassemia, phenylketonuria, Lesch-Nyhan syndrome,  
 CC severe combined immunodeficiency (SCID), cystic fibrosis, Duchenne's  
 CC muscular dystrophy, inherited emphysema, familial hypercholesterolemia,  
 CC diabetes, hypopituitarism, adenine deaminase deficiency, alpha1-  
 CC antitrypsin deficiency, Gaucher's syndrome, anaemia, infections such as  
 CC HIV infection, high blood pressure, Alzheimer's disease, autoimmune or  
 CC inflammatory disease or graft versus host disease. RRV's are capable of  
 CC transfer over prolonged periods of time.  
 CC  
 Db 1481 tctctgaagagaatggtcccaatgagcctgagccactgctccttactactatctt 961  
 CC ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 902 tctctgaagagaatggtcccaatgagcctgagccactgctccttactactatctt 961  
 CC ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 1481 tctctgaagagaatggtcccaatgagcctgagccactgctccttactactatctt 961  
 CC ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 962 ctcagtgagcctgtgtaaaagacttgatcagagcctca 1000  
 CC ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 1541 ctcagtgagcctgtgtaaaagacttgatcagagcctca 1579  
 CC ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 RESULT 11  
 AAV19581  
 ID AAV19581 standard; cDNA; 4832 BP.  
 AC AAV19581;  
 XX  
 DT 06-AUG-1998 (first entry)  
 XX  
 DE Human factor VIII beta-domain deleted SQN deletion cDNA sequence.  
 XX  
 KW Replication defective; recombinant retrovirus; RRV; therapeutic protein;  
 KW haemophilia; thrombosis; hypercoagulable disorder; liver disease; human;  
 KW hepatitis; thalassemia; phenylketonuria; Lesch-Nyhan syndrome; diabetes;  
 KW cystic fibrosis; Duchenne's Muscular Dystrophy; hypercholesterolemia;  
 KW hypopituitarism; adenine deaminase deficiency; HIV infection; anaemia;  
 KW Gaucher's syndrome; high blood pressure; Alzheimer's disease; autoimmune;  
 KW inflammatory disease; factor VIII; ss.  
 KW  
 OS Homo sapiens.  
 XX  
 XX WO9800541-A2.  
 PN 08-JAN-1998.  
 PD 02-JUL-1997; 97WO-US11784.  
 PF 04-JUN-1997; 97US-0869309.  
 PR 03-JUL-1996; 96US-0645601.  
 PR 13-AUG-1996; 96US-0696381.  
 XX  
 XX (CHIR ) CHIRON CORP.  
 PA  
 XX Allen JR, Barber JR, Boder M, Chang SM, Chong K;  
 PI De LA VEGA D, Depoloni J, Greengard J, Hsu DC, Ibanez CE;  
 PI Jolly DJ, Lee R, Mittenstaedt DM, Prussak CE, Respass JG;  
 XX WPI: 1998-086966/08.  
 DR P-PSDB: AAW6246.  
 DR  
 XX  
 PT New replication defective recombinant retro-viruses - which can be  
 PT administered to provide long term systemic expression of therapeutic  
 PT protein in blood, useful in, e.g. treating hyper-coagulable  
 PT disorders  
 XX  
 XX Example 28; Pages 210-213; 272pp; English.  
 XX  
 CC This cDNA encodes the beta-domain deleted SQN deletion protein of human  
 CC factor VIII. This is used in the construction of recombinant retroviral  
 CC vectors expressing human factor VIII. The invention provides the  
 CC preparation of replication defective recombinant retrovirus (RRV)  
 CC expressing a therapeutic protein. The RRV preparation is resistant to  
 CC degradation by human complement and is capable of inducing long term  
 CC systemic expression of the therapeutic protein when administered  
 CC intravenously to a human. The long term systemic expression results in a  
 CC measurable level of the therapeutic protein being produced in the blood  
 CC of the human for a period of at least 30 days after the administration of

CC the RRV vector preparation. RRV's can be used for in vivo delivery of  
 CC therapeutic protein to treat, e.g. haemophilia A, haemophilia B,  
 CC thrombosis, hypercoagulable disorders, liver diseases such as hepatitis,  
 CC disorders such as thalassemia, phenylketonuria, Lesch-Nyhan syndrome,  
 CC severe combined immunodeficiency (SCID), cystic fibrosis, Duchenne's  
 CC muscular dystrophy, inherited emphysema, familial hypercholesterolemia,  
 CC diabetes, hypopituitarism, adenine deaminase deficiency, alpha1-  
 CC antitrypsin deficiency, Gaucher's syndrome, anaemia, infections such as  
 CC HIV infection, high blood pressure, Alzheimer's disease, autoimmune or  
 CC inflammatory disease or graft versus host disease. RRV's are capable of  
 CC transfer over prolonged periods of time.  
 CC  
 Db 131 caccagaagaatgagctccacacgtctcttctgagccttggcattgagtc 545  
 CC ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 426 catgcaaatagagctccacacgtctcttctgagccttggcattgagtc 545  
 CC ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 71 catgcaaatagagctccacacgtctcttctgagccttggcattgagtc 545  
 CC ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 486 caccagaagaatgagctccacacgtctcttctgagccttggcattgagtc 545  
 CC ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 131 caccagaagaatgagctccacacgtctcttctgagccttggcattgagtc 545  
 CC ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 546 cggctgagcgcctgctgagcagaatcttcctcctgagtcgaatgcaaatcttccatcaa 605  
 CC ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 191 cggctgagcgcctgctgagcagaatcttcctcctgagtcgaatgcaaatcttccatcaa 605  
 CC ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 606 caccctagtcgctgacaaaagaactctgttctgtagaattcagatcacccttccaacat 665  
 CC ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 251 caccctagtcgctgacaaaagaactctgttctgtagaattcagatcacccttccaacat 665  
 CC ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 666 cggctagcagaagcaccctgagtcgctgtagcctgagtcgaatgcaaatcttccatcaa 725  
 CC ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 311 cggctagcagaagcaccctgagtcgctgtagcctgagtcgaatgcaaatcttccatcaa 725  
 CC ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 726 tgatcagtgatcattacacttaagaacatggtctccacacgtcagtcctcatgctgt 785  
 CC ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 371 tgatcagtgatcattacacttaagaacatggtctccacacgtcagtcctcatgctgt 785  
 CC ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 766 tgggtaccctcagcgaagctcttgaggagagcctggaatgcaaatcttccatcaa 845  
 CC ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 431 tgggtaccctcagcgaagctcttgaggagagcctggaatgcaaatcttccatcaa 845  
 CC ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 846 ggaagaagaagaatgataaagcttccctgctggaagcacaataatgcttgcagagtcct 905  
 CC ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 491 ggaagaagaagaatgataaagcttccctgctggaagcacaataatgcttgcagagtcct 905  
 CC ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 906 gaaagaagaatgagtcacatgagcctcctgagccactgctccttactactatcttctca 965  
 CC ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 551 gaaagaagaatgagtcacatgagcctcctgagccactgctccttactactatcttctca 965  
 CC ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 966 tctgagcctgtgtaaaagacttgatcagagcctca 1000  
 CC ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 611 tctgagcctgtgtaaaagacttgatcagagcctca 645  
 CC ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 RESULT 12  
 AAV15338  
 ID AAV15338 standard; DNA; 4832 BP.  
 AC AAV15338;  
 XX  
 DT 20-JUL-1998 (first entry)  
 XX  
 DE Human factor VIII SQN deletion mutant DNA.  
 KW Factor VIII; blood clotting; haemophilia A; gene therapy;  
 KW retrovirus; vector; human; ss.



CC mammalian cells, or Bacillus or Kluyveromyces microbial cells,  
 CC for prodn. of a Factor-VIII derivative useful in the treatment  
 CC of haemophilia A. The Factor-VIII derivative has the formula  
 CC NR-IR-CR, where NR is amino acids 1-712 of Factor-VIII, IR is a  
 CC short peptide (AAR86864), and CR is amino acids 1638-2332 of  
 CC Factor-VIII.  
 XX  
 SQ Sequence 6300 BP; 2026 A; 1340 C; 1323 G; 1611 T; 0 other;

Query Match 57.5%; Score 575; DB 17; Length 6300;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-155;  
 Matches 575; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 426 catgcaataagaactctccacgtctctctctgctcttgcctttgcgattcgttaagc 485  
 Db 6 catgcaataagaactctccacgtctctctctgctcttgcctttgcgattcgttaagc 485  
 QY 486 caccagaataatacctctggtgagtgagtgatgcatggagataatgcaagtgatct 65  
 Db 66 caccagaataatacctctggtgagtgagtgatgcatggagataatgcaagtgatct 65  
 QY 546 cgtgagctgctctggtgagtgagtgatgcatggagataatgcaagtgatct 125  
 Db 126 cgtgagctgctctggtgagtgagtgatgcatggagataatgcaagtgatct 125  
 QY 606 cactcaagtcgtgacaaagaagactctgtttgagaattacagagacactttcaaat 185  
 Db 186 cactcaagtcgtgacaaagaagactctgtttgagaattacagagacactttcaaat 185  
 QY 666 cgttaagcgaagcagccctggtgagtgagtgatgcatggagataatgcaagtgatct 245  
 Db 246 cgttaagcgaagcagccctggtgagtgagtgatgcatggagataatgcaagtgatct 245  
 QY 726 tgatacagtggtcattacacttaagaacatggcttccatctgtcacttcaatgctgt 305  
 Db 306 tgatacagtggtcattacacttaagaacatggcttccatctgtcacttcaatgctgt 305  
 QY 786 tgggtatcctactggaagactctggaagtgagtgatgcatggagataatgcaagtgatct 365  
 Db 366 tgggtatcctactggaagactctggaagtgagtgatgcatggagataatgcaagtgatct 365  
 QY 846 ggaagaagaatgataaagctctcctggtgagtgagtgatgcatggagataatgcaagtgatct 425  
 Db 426 ggaagaagaatgataaagctctcctggtgagtgagtgatgcatggagataatgcaagtgatct 425  
 QY 906 gaaagaagaatgataaagctctcctggtgagtgagtgatgcatggagataatgcaagtgatct 485  
 Db 486 gaaagaagaatgataaagctctcctggtgagtgagtgatgcatggagataatgcaagtgatct 485  
 QY 966 tgggtatcctactggaagactctggaagtgagtgatgcatggagataatgcaagtgatct 1000  
 Db 546 tgggtatcctactggaagactctggaagtgagtgatgcatggagataatgcaagtgatct 580

RESULT 14  
 ID AAN81439  
 XX AAN81439 standard; cDNA; 8241 BP.  
 AC AAN81439;  
 DT 08-OCT-1990 (first entry)  
 XX  
 DE Factor VIII cDNA insert of pCLB89.  
 XX  
 KW Factor VIII; blood clotting disorders; ss.  
 OS Homo sapiens.  
 XX  
 XX  
 FH Key Location/Qualifiers  
 FT 7..7062  
 FT /tag= a  
 FT CDS /product=Factor VIII

FT sig-peptide 7..63  
 FT mat-peptide /tag= b  
 FT 64..7059  
 XX /tag= c  
 PN EP253455-A.  
 XX  
 PD 20-JAN-1988.  
 XX  
 PE 20-JUL-1987; 87EP-0201379.  
 XX  
 PR 18-JUL-1986; 86GB-0017594.  
 XX  
 PR 03-DEC-1986; 86GB-0030659.  
 XX  
 PA (KONN) GIST-BROCADES NV.  
 XX  
 PI Van Ooyen AJJ, Andreoli PM, Van Mourik JA, Pannekoek H;  
 XX  
 DR WPI: 1988-015822/03.  
 DR  
 DR P-PSDB; AAP81113.

PT Expression system for producing factor VIII polypeptide in  
 PT microbial cells - contg. open reading frame and functional  
 PT initiation and termination regions.  
 XX  
 PS Disclosure; 02pp; English.  
 CC  
 CC The cDNA was prepd. from human liver RNA. Clone pCLB89 was  
 CC produced by combining inserts from several positive clones  
 CC isolated from the cDNA library. The sequence can be inserted  
 CC into expression vectors for the prodn. of recombinant factor  
 CC VIII. The protein produced is useful for treating blood clotting  
 CC disorders.  
 XX  
 SQ Sequence 8241 BP; 2625 A; 1759 C; 1703 G; 2154 T; 0 other;

Query Match 57.5%; Score 575; DB 9; Length 8241;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-155;  
 Matches 575; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 426 catgcaataagaactctccacgtctctctctgctcttgcctttgcgattcgttaagc 485  
 Db 6 catgcaataagaactctccacgtctctctctgctcttgcctttgcgattcgttaagc 485  
 QY 486 caccagaataatacctctggtgagtgagtgatgcatggagataatgcaagtgatct 65  
 Db 66 caccagaataatacctctggtgagtgagtgatgcatggagataatgcaagtgatct 65  
 QY 546 cgtgagctgctctggtgagtgagtgatgcatggagataatgcaagtgatct 125  
 Db 126 cgtgagctgctctggtgagtgagtgatgcatggagataatgcaagtgatct 125  
 QY 606 cactcaagtcgtgacaaagaagactctgtttgagaattacagagacactttcaaat 185  
 Db 186 cactcaagtcgtgacaaagaagactctgtttgagaattacagagacactttcaaat 185  
 QY 666 cgttaagcgaagcagccctggtgagtgagtgatgcatggagataatgcaagtgatct 245  
 Db 246 cgttaagcgaagcagccctggtgagtgagtgatgcatggagataatgcaagtgatct 245  
 QY 726 tgatacagtggtcattacacttaagaacatggcttccatctgtcacttcaatgctgt 305  
 Db 306 tgatacagtggtcattacacttaagaacatggcttccatctgtcacttcaatgctgt 305  
 QY 786 tgggtatcctactggaagactctggaagtgagtgatgcatggagataatgcaagtgatct 365  
 Db 366 tgggtatcctactggaagactctggaagtgagtgatgcatggagataatgcaagtgatct 365  
 QY 846 ggaagaagaatgataaagctctcctggtgagtgagtgatgcatggagataatgcaagtgatct 425  
 Db 426 ggaagaagaatgataaagctctcctggtgagtgagtgatgcatggagataatgcaagtgatct 485



QY 906 gaagagaatggtccatggtcctgacccactggtccttactactatattcttca 965  
 |||||||  
 Db 486 gaagagaatggtccatggtcctgacccactggtccttactactatattcttca 545  
 |||||||  
 QY 966 tgtgacctgtgtaaaagacttgaaatcaggtccta 1000  
 |||||||  
 Db 546 tgtgacctgtgtaaaagacttgaaatcaggtccta 580  
 |||||||

## RESULT 15

AAAB1096 standard; cDNA: 8241 BP.

AAAB1096:

29-OCT-1990 (first entry)

XX cDNA sequence encoding human Factor VIII signal sequence and mature  
 DE protein on plasmid pCLB89.

XX Haemophilia A; blood clotting; treatment; diagnosis; ss.

XX Homo sapiens.

XX Location/Qualifiers

XX Key 7..63

XX CDS /tag= a

XX /product=Signal peptide

XX mat\_peptide 64..7062

XX /tag= b

XX EP294910-A.

XX 14-DEC-1988.

XX 13-JUN-1988; 88EP-0201209.

XX 12-JUN-1987; 87EP-0201121.

XX 03-JUN-1988; 88EP-0201209.

XX (KONN) GIST-BROCADES NV.

XX van Ooyen AJJ, Pannekoek H, Verbeet MP, van Leen RW.

XX WPI: 1988-355361/50.

XX P-PSDB: AAP80659.

XX Proteins having Factor VIII activity

XX comprising deletion mutant proteins of Factor VIII in which

XX central region has been deleted

XX Example: Fig 1: 37pp; English:

XX The isolation of Factor VIII mRNA from human liver, and the preparation,

XX purification and identification of cDNA and its assembly in the plasmid

XX pPE121 resulting in plasmid pCLB89 have been described in patent

XX application EP 0253455. The Factor VIII polypeptides of the invention

XX include deletion mutant proteins of Factor VIII in which the central

XX region or "B domain" as well as a portion of the 92 kD region has been

XX deleted. The polypeptides have enhanced Factor VIII activity and/or

XX decreased immunogenicity and can be used for the treatment of

XX Haemophilia A. They can also be used to prepare antibodies. The Abs.

XX can be used to determine the concn. of Factor VIII in a body fluid.

XX Sequence 8241 BP; 2625 A; 1759 C; 1703 G; 2154 T; 0 other;

Query Match 57.5%; Score 575; DB 9; Length 8241;

Best Local Similarity 100.0%; Pred. No. 2.4e-155; Indels 0; Gaps 0;

Matches 575; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 6 catgcaaatagagctctccacctgctcttcttctgaccttggagatctgctttagtgc 65  
 QY 486 caccagaagatatactactggtggtcagtgagactgtcatggaactatagcaagtgtatc 545  
 |||||||  
 Db 66 caccagaagatatactactggtggtcagtgagactgtcatggaactatagcaagtgtatc 125  
 |||||||  
 QY 546 cgtgagctgctg tggagcgaagattctctctctctctctctctctctctctctctctct 605  
 |||||||  
 Db 126 cgtgagctgctg tggagcgaagattctctctctctctctctctctctctctctctctct 185  
 |||||||  
 QY 606 cactcagtgctggtgtaaaagactctgtttagaattcaggaatcaggaatcaggaatcaggaat 665  
 |||||||  
 Db 186 cactcagtgctggtgtaaaagactctgtttagaattcaggaatcaggaatcaggaatcaggaat 245  
 |||||||  
 QY 666 cgttaagccaagcacccttgagtggtctgtctgagtcctaccatccatccaggtgaggttta 725  
 |||||||  
 Db 246 cgttaagccaagcacccttgagtggtctgtctgagtcctaccatccatccaggtgaggttta 305  
 |||||||  
 QY 726 tgatacagtggtcatatataccttaagaagactggtccatccatccatccatccatccatccat 785  
 |||||||  
 Db 306 tgatacagtggtcatatataccttaagaagactggtccatccatccatccatccatccatccat 365  
 |||||||  
 QY 786 tgggtatcctactggaagctctctgaggaagctggaatgataatgataatgataatgataatgata 845  
 |||||||  
 Db 366 tgggtatcctactggaagctctctgaggaagctggaatgataatgataatgataatgataatgata 425  
 |||||||  
 QY 846 ggaagaagaagatataaagctctctgaggaagctggaatgataatgataatgataatgataatgata 905  
 |||||||  
 Db 426 ggaagaagaagatataaagctctctgaggaagctggaatgataatgataatgataatgataatgata 485  
 |||||||  
 QY 906 gaaagaagatggtccatggtcctgagccactggtccttactactatattcttca 965  
 |||||||  
 Db 486 gaaagaagatggtccatggtcctgagccactggtccttactactatattcttca 545  
 |||||||  
 QY 966 tgtgacctgtgtaaaagacttgaaatcaggtccta 1000  
 |||||||  
 Db 546 tgtgacctgtgtaaaagacttgaaatcaggtccta 580  
 |||||||

Search completed: January 8, 2002, 17:26:36  
 Job time: 12576 sec





```

1 TYPE: DNA
2 ORGANISM: Artificial Sequence
3 FEATURE:
4 OTHER INFORMATION: Description of Artificial Sequence: synthetic
5 US-09-364,862-13

```

Query Match	100.0%	Score 1000;	DB 4;	Length 11933;
Best Local Similarity	100.0%;	Pred. No. 5.2e-308;		
Matches 1000;	Conservative	0;	Mismatches	0

[illegible]



```

? ADDRESS: Foley & Lardner
? STREET: 3000 K Street, N.W., Suite 500
? City: Washington
? STATE: D.C.
? COUNTRY: USA
? ZIP: 20007-5109
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentin Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/882,083
? FILING DATE:
? CLASSIFICATION: 514
? PRIORITY APPLICATION DATA:
? APPLICATION NUMBER: US 08/558,107
? FILING DATE: 13-NOV-1995
? ATTORNEY/AGENT INFORMATION:
? NAME: ISACSON, John P.
? REGISTRATION NUMBER: 33,715
? REFERENCE/DOCKET NUMBER: 30472/212
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (202)672-5300
? TELEFAX: (202)672-5399
? TELEX: 904136
? INFORMATION FOR SEQ ID NO: 1:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 5035 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 35..5017
? US-08-882-083-1

```

Query Match 57.7%; Score 577.4; DB 2; Length 5035;  
Best Local Similarity 99.8%; Pred. No. 1.2e-173;  
Matches 578; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Dy 422 ccacccatgcaaatagagcctccaccctgtcttcttgctgcctttggatcttcgttta 481  
Db 30 CCACCATGTGAAATAAGAGCTTCACACTGGTCTTCTTGCTGCGCTTTTGCAATCTCTTTA 89  
Dy 482 gtgccaccgaagaatactaccctcggtgtgcagtggaactgtcatbggaactatatgcaaa 541  
Db 90 GTGCCACCAGAAGATATCTACTGCTGGGTGCAGTGAACCTGTCAATGAGCATATATGCAAA 149  
Dy 542 atctcggtgagctgtccctgtgagcgcgaagaattccctccctagagtgcgaaaacttt 601  
Db 150 ATCTCGGTGAGCTGCTGCTGTGGCAGCAAGATTTCCTCTTAGAGTGCACAAAATCTTTT 209  
Dy 602 tcaaacacctcaactgttgtacaanaaagactcgttgttaagaattcaaggatccaacttt 661  
Db 210 TCAAACACCTCAAGCTGCTGTACAANAAGACTCTGTTTTGTAGAATTCAACGATCACCTTTT 269  
Dy 662 acatgcctcaagcgaagccacccttgatyggtctgcttagtgcctcacatccaggctgag 721  
Db 270 ACATGCTTAACCCAAAGGCCACCTTGATGGGTCTGTAGTCTTACCATTCAGGCTTAGG 329  
Dy 722 ttatlgatacagtggtgatcttacacttaagaacatgagttcccactcctgtaagttcata 781  
Db 330 TTATGTATAGTAGTGTGTAATTAACCTTAAGAACATGGCTTCCCATCTGTCAGTCTTCATG 389  
Dy 782 ctgtgtgtgtatcctctaactggaagactctgaaggagctgtaatalgatatgatacagcag 841  
Db 390 CTGTGTGTGTATCTTACTTGTGAAGAACTTCTGAGGAGAGCTGAATATGATGATCAACCA 449  
Dy 842 aaagggaagaagaatgataaagcttccctcgtgtggaagccataatatgtctggcagg 901



[illegible]

RESULT 8  
 5171844-1  
 : Patent No. 5171844  
 : APPLICANT: VAN OOYEN, ALBERT J.J.; PANNEKOEK, HANS; VERBEET,  
 : MARTINUS P.; VAN LEEN, ROBERT W.  
 : TITLE OF INVENTION: PROTEINS WITH FACTOR VIII ACTIVITY  
 : PROCESS FOR THEIR PREPARATION USING GENETICALLY-ENGINEERED CELLS  
 : AND PHARMACEUTICAL COMPOSITIONS CONTAINING THEM  
 : NUMBER OF SEQUENCES: 12  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/07/205,226  
 : FILING DATE: 10-JUN-1988  
 : SEQ ID NO.:1  
 : LENGTH: 8241  
 5171844-1

Query Match	57.5%;	Score 575;	DB 6;	Length 8241,
Best Local Similarity	100.0%;	Pred. NO.	9.2e-173;	
Matches 575; Conservative	0.	Mismatch	0.	

	QY	426	catgcaatagagctctccactgctcttctctgctcttgcatctgccttagtcg	485
QY				
Db	6	catgcaatagagctctccactgctctctctgctcttgcatctgccttagtcg	65	
QY	486	caccagaagatctactcctggtgcatggaacttcattgagacatattgaaagtatct	545	
Db	66	caccagaagatctactcctggtgcatggaacttcattgagacatattgaaagtatct	125	

QY	546	cgtgtgacctgctgtgtgacgcaaaattctctccctcagagtgccaattcttccattcaa	605
Db	126	cgtgtgacctgctgtgtgacgcaaaattctctccctcagagtgccaattcttccattcaa	185
QY	606	caactcagctcagttgtacaaaagaactctgtttgtagaattcaaggatcaatttcaaat	665
Db	186	caactcagctcagttgtacaaaagaactctgtttgtagaattcaaggatcaatttcaaat	245
QY	666	cgttaagccaaagccaaacctgtatgtgtctcgtacgttccaccatccagctgaagttta	725
Db	246	cgttaagccaaagccaaacctgtatgtgtctcgtacgttccaccatccagctgaagttta	305
QY	726	tgtatcagtggtcattatcaacttaagaanaatggtctcccatcctgtcagttcattcgtt	785
Db	306	tgtatcagtggtcattatcaacttaagaanaatggtctcccatcctgtcagttcattcgtt	365
QY	786	tgtgttatctcactctggaagctctctctgaggagcttgaaatgatgatccaaagccaag	845
Db	366	tgtgttatctcactctggaagctctctctgaggagcttgaaatgatgatccaaagccaag	425
QY	846	ggagaagaagaatgataaagaactctccctctgtgtgaaagcatatcatatgtctggcaggtct	905
Db	426	ggagaagaagaatgataaagaactctccctctgtgtgaaagcatatcatatgtctggcaggtct	485
QY	906	gaaagaagaatggtccaaatggtcctctgaccacatgtgccttatcctactatattcttctca	965
Db	486	gaaagaagaatggtccaaatggtcctctgaccacatgtgccttatcctactatattcttctca	545
QY	966	tgtgacctgtgtaaagaacttgaaatcagagctcca	1000
Db	546	tgtgacctgtgtaaagaacttgaaatcagagctcca	580

RESULT 9  
US-08-366-851A-1  
; Sequence 1, Application US/08366851A  
; Patent No. 5681746  
; GENERAL INFORMATION:

REFERENCE: Bodmer, Mordechai  
 APPLICANT: De Polo, Nicolas J.  
 APPLICANT: Hsu, David Chi-Tang  
 APPLICANT: Chang, Steven  
 TITLE OF INVENTION: Retroviral Delivery of Full Length Factor VIII  
 NUMBER OF SEQUENCES: 3  
 CORRESPONDENCE ADDRESS:

ADDRESS: Viagene, Inc.  
STREET: 11055 Roselle Street  
CITY: San Diego  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 92121

COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/366,851A  
FILING DATE:

CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Chambers, Daniel M.  
REGISTRATION NUMBER: 34,561  
REFERENCE/DOCKET NUMBER: 930049.438  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 452-1288  
TELEFAX: (619) 452-2616  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8667 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: both



[illegible]

OY 306 gaaggaagaatggtcccaatgagctctgacccactgtgaccttaactactatattcttca 965  
 Dd 630 GAAAGAGANTGGTCCCAATGGCCCTCGACCCACGTGCTTACCTCATATCTTTCTCA  
 OY 966 ttgtggacccgtgtaaaagacttgaattcaagccctca 1000  
 Dd 690 TTGTGGACCTGTTAAAGACTTGATTTGATTCAGGCCCTTCA 724

RESULT 11  
US-08-251-

```

1 Sequence 3, Application US/08251937A
2 Patent No. 5583209
3 GENERAL INFORMATION:
4 APPLICANT: Lollar, John S.
5 APPLICANT: Runge, Marshall S.
6 TITLE OF INVENTION: Hybrid Human/Porcine Factor VIII
7 NUMBER OF SEQUENCES: 10
8 CORRESPONDENCE ADDRESS:
9 ADDRESSEE: Kildpatrick & Cody
10 STREET: 1100 Peachtree Street
11 CITY: Atlanta
12 STATE: Georgia
13 COUNTRY: US
14 ZIP: 30309
15 COMPUTER READABLE FORM:
16 MEDIUM TYPE: Floppy disk
17 COMPUTER: IBM PC compatible
18 OPERATING SYSTEM: PC-DOS/MS-DOS
19 SOFTWARE: Patent In Release #1.0, Version #1.25
20 CURRENT APPLICATION DATA:
21 APPLICATION NUMBER: US/08/251,937A
22 FILING DATE: 31-MAY-1994
23 CLASSIFICATION: 435
24 PRIOR APPLICATION DATA:
25 APPLICATION NUMBER: US 07/864,004
26 FILING DATE: 07-APR-1992
27 ATTORNEY/AGENT INFORMATION:
28 NAME: Pratt, John S.
29 REGISTRATION NUMBER: 29,476
30 REFERENCE/DOCKET NUMBER: EBU106DIV
31 TELECOMMUNICATION INFORMATION:
32 TELEPHONE: 404-815-6367
33 TELEFAX: 404-815-6555
34 INFORMATION FOR SEQ ID NO: 3:
35 SEQUENCE CHARACTERISTICS:
36 LENGTH: 9009 base pairs
37 TYPE: nucleic acid
38 STRANDEDNESS: single
39 TOPOLOGY: linear
40 MOLECULE TYPE: cDNA
41 HYPOTHETICAL: NO
42 ANTI-SENSE: NO
43 FRAGMENT TYPE: N-terminal
44 ORIGINAL SOURCE:
45 ORGANISM: Homo sapien
46 TISSUE TYPE: Liver
47 FEATURE:
48 NAME/KEY: misc_feature (Domain Structure)
49 LOCATION: 5001 .. 7053
50 OTHER INFORMATION: /note="Equivalent to the A3-C1-C2
51 OTHER INFORMATION: domain"
52 NAME/KEY: misc_feature (Domain Structure)
53 LOCATION: 1 .. 2277
54 OTHER INFORMATION: /note="Equivalent to the A1-A2
55 OTHER INFORMATION: domain"
56 US-08-251-937A-3

```

Query Match	57.38;	Score 573.4;	DB 1;	Length 9009;
Best Local Similarity	99.48;	Pred. No. 3.1e-172;		

	Matches 574;	Conservative	0;	Mismatches	1;	Indels	0;	Gaps	0;
QY	426	catgcaaatagagctctccacctgctcttctctgctcttgccttctgcatctgctctttagc							485
Db	150	CATGCAAAATAGAGCTCTCCACCTGCTTTTCTCTGTGCTTTTGGATTTGCTTTAGTGC							209
QY	486	caccaaaatactactactcctgggtgcagtgaaactgcatgggactatatagaagtgatct							545
Db	210	CACCGAAGATATCAACCTGGGTGCAGTGAACTGTACTGGGACTATATATCAAGAAGATCT							269
QY	546	cggtagactgctctgtagcgcgaatatctctctctctctctctctctctctctctctctct							605
Db	270	CGGTGAGCTGGCTGTGGACGCAAGATTTCTCTCTTAAGTCCAAATCTTTTCCATTGAA							329
QY	606	caccctcagctgctgtacaaaaaagactctgtcttgtagattcagaaggtctaccttcaacat							665
Db	330	CACCTCAGTCTGTACAAAAAGACTCTGTTTGTAGATTCACGGTTCACCTTTTCAACAT							389
QY	666	cgttaagccaagccacccctggaatggctctgctctgctctctctctctctctctctctct							725
Db	390	CGCTAAGCCAAAGCCACCTCGATGGGCTGTGTAGTCTCTACATCCAGGCTGAGGTTTA							449
QY	726	tgatacagtggtcatatcatacttaagaacaatggctctccatctctgctcagttcctatgctgt							785
Db	450	TGATACAGTGCATATPACACTTAAGAACATGGCTTCCATCTCTGTACGTCTTATCTCTGT							509
QY	786	tggctgctctacttgaaagctctcgaaaggactgaaatgatatcatagacacagtgcaag							845
Db	510	TGGTGATCTCTACTGAAAAGCTTGTGAGGAGGCTGATATGATGATCAGACCAAGTAAAG							569
QY	846	ggaagaagaaga tga taaagctctccctgggtgggaagcaataatgctcggcaggtcct							905
Db	570	GGAGAAAAGAAGTGTATAAGTCTTCCCTGTGGAGGCCATCATATATGTGTGCAGAGTCT							629
QY	906	gaagaagaatggtccaatagcctctgaaccactgtgcttactactatcatatcttctca							965
Db	630	GAAGAAGATGTGCTCAATGGCTCTGTGACCCACATGTGCTTACTACTATATCTTTTCYA							689
QY	966	tgtgacctgtgtaaaagactgaaatcagaagctca 1000							
Db	690	TGTGACCTGTGTAAGAAGCTTGAATTCAGAGCCCTCA 724							

RESULT 12  
US-08-212-

Sequence 1, Application US/08212133A  
Patent No. 5663060  
GENERAL INFORMATION:  
APPLICANT: Iollet, John S.  
APPLICANT: Runge, Marschall S.  
TITLE OF INVENTION: Hybrid Human/Animal Factor VII  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Kilpatrick & Cody  
STREET: 100 Peachtree Street  
CITY: Atlanta  
STATE: Georgia  
COUNTRY: US  
ZIP: 30303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/212,133A  
FILING DATE: March 11, 1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/864,004  
FILING DATE: 07-ARR-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Pabst, Patricia L.

Wed Jan 9 14:38:20 2002

us-09-740-211-13-copy\_1\_1000.rn1

Page 9

REGISTRATION NUMBER: 31,284  
REFERENCE/DOCKET NUMBER: EMU/76677  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 404-572-6508  
TELEFAX: 404-572-6555  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9009 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Homo sapien  
TISSUE TYPE: Liver  
FEATURE:  
NAME/KEY: misc.feature (Domain Structure)  
LOCATION: 5125..7053  
OTHER INFORMATION: /note="Equivalent to the A3-C1-C2  
OTHER INFORMATION: domain"  
FEATURE:  
NAME/KEY: misc.feature (Domain Structure)  
LOCATION: 1..2277  
OTHER INFORMATION: /note="Equivalent to the A1-A2 domain."  
FEATURE:  
NAME/KEY: Domain  
LOCATION: 1..2277  
OTHER INFORMATION: /note="CDNA encoding human factor  
OTHER INFORMATION: VIII."  
US-08-212-133A-1

Query Match 57.3%; Score 573.4; DB 1; Length 9009;  
Best Local Similarity 99.8%; Pred. No. 3.1e-172;  
Matches 574; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

426 catgcaatagagctccaccgctcttctgctgcttgcgattcgtttagtc 485  
150 CATGCAATAGAGCTCCACCCTCTTCTGCTTTCGCTTTTCGATTGCTG 209  
486 caccagaagatactaccgagtgagtgagtgagtgagtgagtgagtgag 545  
210 CACCAGAAGATACTACCAGTGAGTGAGTGAGTGAGTGAGTGAGTGAG 269  
546 cagtgagctgctgagtgagtgagtgagtgagtgagtgagtgagtgag 605  
270 CGTGAGCTGCTGAGTGAGTGAGTGAGTGAGTGAGTGAGTGAGTGAG 329  
606 cactcagctgctgagtgagtgagtgagtgagtgagtgagtgagtgag 665  
330 CACTCAGCTGCTGAGTGAGTGAGTGAGTGAGTGAGTGAGTGAGTGAG 389  
666 cgttaagcaagagcagctgagtgagtgagtgagtgagtgagtgagtg 725  
390 CGCTAAGCAAGAGCCACCCTGAGTGAGTGAGTGAGTGAGTGAGTGAG 449  
726 tgatacagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 785  
450 TGATACAGTGAGTGAGTGAGTGAGTGAGTGAGTGAGTGAGTGAGTGAG 509  
786 tgggtatctctgagtgagtgagtgagtgagtgagtgagtgagtgagtg 845  
510 TGGTGATCTCTGAGTGAGTGAGTGAGTGAGTGAGTGAGTGAGTGAG 569  
846 gggagaagaagatgagtgagtgagtgagtgagtgagtgagtgagtgag 629  
570 GGAGAAGAAGATGAGTGAGTGAGTGAGTGAGTGAGTGAGTGAGTGAG 629  
906 gaaagaagaatgagtgagtgagtgagtgagtgagtgagtgagtgagtg 965  
630 GAAAGAAGAATGAGTGAGTGAGTGAGTGAGTGAGTGAGTGAGTGAGTG 689

QY 966 tgggagctgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1000  
Db 690 TGGGAGCTGAGTGAGTGAGTGAGTGAGTGAGTGAGTGAGTGAGTG 724

RESULT 13  
US-08-474-503-1  
Sequence 1, Application US/08474503  
Patent No. 574446  
GENERAL INFORMATION:  
APPLICANT: Emory University  
TITLE OF INVENTION: Hybrid Human/Animal Factor VIII  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Kilpatrick & Cody  
STREET: 1100 Peachtree Street, Suite 2800  
CITY: Atlanta  
STATE: Georgia  
COUNTRY: US  
ZIP: 30309  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/474,503  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Pratt, John S.  
REGISTRATION NUMBER: 29,476  
REFERENCE/DOCKET NUMBER: EMU106CIP(3)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 404-815-6500  
TELEFAX: 404-815-6555  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9009 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Homo sapien  
TISSUE TYPE: Liver  
FEATURE:  
NAME/KEY: misc.feature (Domain Structure)  
LOCATION: 5125..7053  
OTHER INFORMATION: /note="Equivalent to the A3-C1-C2  
OTHER INFORMATION: domain"  
FEATURE:  
NAME/KEY: misc.feature (Domain Structure)  
LOCATION: 1..2277  
OTHER INFORMATION: /note="Equivalent to the A1-A2 domain."  
FEATURE:  
NAME/KEY: Domain  
LOCATION: 1..2277  
OTHER INFORMATION: /note="CDNA encoding human factor  
OTHER INFORMATION: VIII."  
US-08-474-503-1

Query Match 57.3%; Score 573.4; DB 1; Length 9009;  
Best Local Similarity 99.8%; Pred. No. 3.1e-172;  
Matches 574; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

426 catgcaatagagctccaccgctcttctgctgcttgcgattcgtttagtc 485  
150 CATGCAATAGAGCTCCACCCTCTTCTGCTTTCGCTTTTCGATTGCTG 209



Db 630 GAAAGAGATGTCAGATGACCTGACCCACTGTCCTTACTACTGATATCTTCTCA 689

Qy 966 tctgacctgtgtaaaagacttgatcaggcctca 1000

Db 690 TGtGACCTGCTAAAGACTTGAATTCAGGCTCA 724

## RESULT 15

US-09-037-601-1

Sequence 1, Application US/09037601

Patent No. 6180371

GENERAL INFORMATION:

APPLICANT: Lohat, John S.

TITLE OF INVENTION: Hybrid Human/Animal Factor VIII

NUMBER OF SEQUENCES: 40

CORRESPONDENCE ADDRESS:

ADDRESSEE: Greenlee, Winner and Sullivan, P.C.

STREET: 5370 Manhattan Circle, Suite 201

CITY: Boulder

STATE: Colorado

COUNTRY: USA

ZIP: 80303

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/037, 601

FILING DATE: 26-JUN-1996

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO PCT/US94/13200

FILING DATE: 15-NOV-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/212,133

FILING DATE: 11-MAR-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/864,004

FILING DATE: 07-APR-1992

ATTORNEY/AGENT INFORMATION:

NAME: Ferber, Donna M.

REGISTRATION NUMBER: 33,878

REFERENCE/DOCKET NUMBER: 75-95F

TELEPHONE: 303/499-8089

TELEFAX: 303/499-8089

INFORMATION FOR SEQ. ID NO. 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 9009 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: not relevant

MOLECULE TYPE: cDNA to mRNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Homo sapiens

TISSUE TYPE: Liver

FEATURE:

NAME/KEY: misc.feature

LOCATION: 5125..7053

OTHER INFORMATION: /product= "Domain Structure"

OTHER INFORMATION: /note= "Equivalent to the A3-C1-C2 domain"

OTHER INFORMATION: /note= "cDNA encoding human factor-VIII"

Query Match 57.3%; Score 573.4; DB 4; Length 9009;  
Best Local Similarity 99.8%; Pred. No. 3,1e-172;  
Matches 574; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 426 catgcaaatagagctcaccacccctgctcttctgtgctcttggcgatctgctttagtgc 485

Db 150 CATCAATATAGACTCTCCACCTGCTTCTTCTGTCCTTTGCGATTCCTTACGTC 209

Qy 486 caccagaagatctactctggtgacgtgacgtgacgtgacgtgacgtgacgtgacgt 545

Db 210 CACCAAGATGACTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 269

Qy 546 cgtgacgtcctgctgagcgaagattctctctgacgtgacgtgacgtgacgtgacgtgac 605

Db 270 CGGTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 329

Qy 606 caactcagctgctgacgaagactctgcttctgacgtgacgtgacgtgacgtgacgtgac 665

Db 330 CACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 389

Qy 666 cgtlaagcgaagcgaagcgaagcgaagcgaagcgaagcgaagcgaagcgaagcgaagcga 725

Db 390 CGCTAAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 449

Qy 726 tgatcagctgctcattctacgttgaagacgtgacgtgacgtgacgtgacgtgacgtgac 785

Db 450 TGATCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 509

Qy 786 tggatcctcctcctggaagcttctggaagcgtggaatgataatgataatgataatgata 845

Db 510 TGCTGATCTCTCTGGAAGCTTCTGAGGAGCTGGAATGATGATGATGATGATGATGAT 569

Qy 846 ggaagaagaagatgataaagcttctcctggtggaagcgaagcgaagcgaagcgaagcga 905

Db 570 GGAGAAAGAGATGATTAAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 629

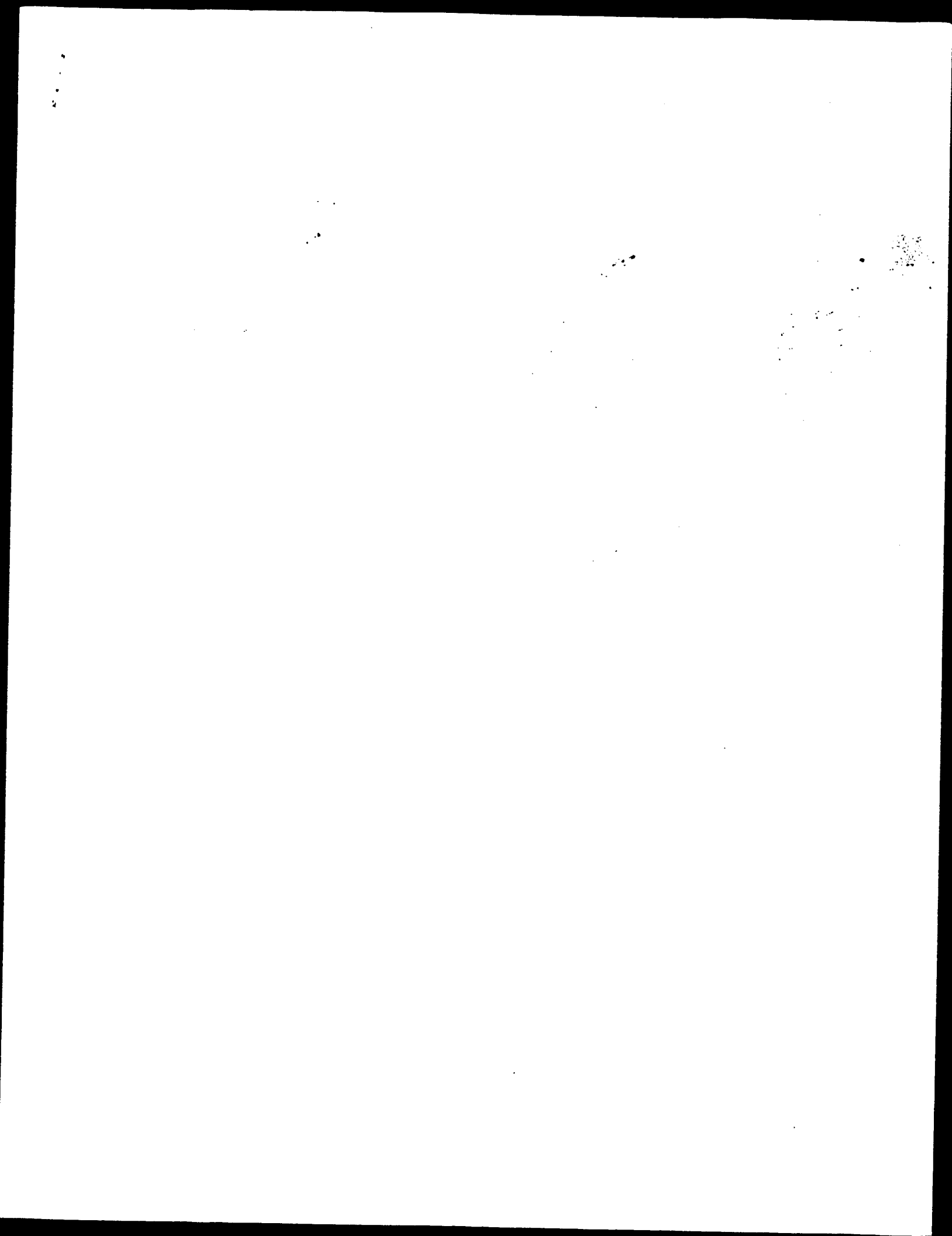
Qy 906 gaaagaagaatgctcactgctcctgacccactgctcctcctcctcctcctcctcctcct 965

Db 630 GAAAGAGATGCTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 689

Qy 966 tctgacctgtgtaaaagacttgatcaggcctca 1000

Db 690 TGtGACCTGCTAAAGACTTGAATTCAGGCTCA 724

Search completed: January 8, 2002, 17:19:07  
Job time: 12198 sec





/note="Organ: skin; Vector: pOTB7; Site-1: XhoI; Site-2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAGGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 233 a 276 c 232 g 212 t 1 others

Query Match 31.5%; Score 315; DB 11; Length 954;  
Best Local Similarity 89.0%; Pred. No. 2,7e-82;  
Matches 421; Conservative 0; Mismatches 40; Indels 12; Gaps 7;

530 atatacaagatgacgtcgtagagctgctgagcgaagaatttccctcagaagtcgca 569  
108 AGATGACACCGAGATGCGGCTGCGAAGAGCGTCCGCGGATTTCTCTAGAGTCCAA 167  
590 aatcttccatcaacacacacacagtcgctgtaaaaaaagactctgtttagaattcag 649  
168 AATCTTTTCCATCAACACCTCACTCGTACAAAAGACTCTGTTGTGATTCACGG 227  
650 atcacccttcaacatcgctgaagcgaagccacccctgtagctctgtagctaccac 709  
228 ATCACCCTTTTACACATGCTTAAGCCAAAGCCACCTGGATGGCTGTGCTTACCA 287  
710 tccagctgaagttatgatagacagtggtacataccttaagaacatgagctccaccct 769  
288 TCCAGGCTGAGGTTTATGATACACTGCTCATTTACCTAGAACATGAGCTTCCATCTG 347  
770 tcaatctcaatgctgtgtgtatcctcactcgtgaagcttctga-aggagctgaatatga 827  
348 TCACTCTTCATGCTGTGTGTATCTTACATGGAAGCTTCTGACGCGACGTCGAATATGA 407  
828 tgatcagacacagtcgca-aggagaaagaagatgata-aagcttccctcgtggaagcca 884  
408 TGATCAGACACAGTCGAAGGAGAGACGATGATCAAGTCTTCCCTGGTGGAGGCCA 467  
885 taact-atgtctggaagctcgt---aaagaagaatgctcaatgagctctgagccactg 939  
468 TACCTCATGCTGTGCGACGCTGCAAAAGACATGCTCCACCTGGCTTGCACCCACATGA 527  
940 -tgctcactactacatattcttcattgagacccgtg-taaagagcttgaat 990  
528 TGCCTACTACTACTCATATCTTCTCATGTGACACTGGCTTACAGACCTTGAAAT 580

RESULT 2  
AA817854/c 456 bp mRNA EST 03-JUL-1999  
LOCUS  
DEFINITION  
UI-R-A0-ae-g-10-0-UI-51 UI-R-A0 Rattus norvegicus cDNA clone  
UI-R-A0-ae-g-10-0-UI 3' similar to gb|U38691|RATCERU Rat norvegicus  
ceruloplasmic mRNA, complete cds, mRNA sequence.  
ACCESSION  
AA817854  
VERSION  
AA817854.1 GI:4198380  
KEYWORDS  
SOURCE  
ORGANISM  
Rattus norvegicus  
Norway rat.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
1 (bases 1 to 456)  
Bonaldo, M.F., Lennon, G. and Soares, M.B.  
Normalization and subtraction: two approaches to facilitate gene  
discovery  
Genome Res. 6 (9), 791-806. (1996)

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
COMMENT  
On Mar 9, 1998 this sequence version replaced gi:2946779.  
Contact: Soares, MB  
Program for Rat Gene Discovery and Mapping  
University of Iowa

451 Eckstein Medical Research Building Iowa City, IA 52242, USA  
Tel: 319 335 8280  
Fax: 319 335 9565  
Email: mscares@iue.weeg.uiowa.edu  
The sequence tag present in the cDNA between the NotI site and the  
oligo-dT track served to identify it as a clone from the normalized  
adult Ovary library. cDNA Library Preparation: M. Fatima Bonaldo,  
Ph.D. Clone distribution: clones will be available through Research  
Genetics. This clone is also available through the I.M.A.G.E.  
Consortium at LUNL (info@image.llnl.gov). IMAGE ID=1767329  
Seq primer: M13 Forward  
POLYA-No.

FEATURES  
source  
Location/Qualifiers  
1..456

/organism="Rattus norvegicus"  
/strain="Sprague-Dawley"  
/db\_xref="taxon:10116"  
/clone="UI-R-A0-ae-g-10-0-UI"  
/clone\_1ib="UI-R-A0"  
/dev\_stage="adult"  
/lab\_host="DH10B (Life Technologies)"  
/note="Vector: pRT3D-Pac (Pharmacia) with a modified  
polylinker. Site-1: Not I; Site-2: Eco RI; This library  
consists of a mixture of individually tagged normalized  
libraries constructed from rat placenta, adult lung, brain  
, liver, kidney, heart, spleen, ovary and muscle. The tag  
is a string of 3-5 nucleotides present between the Not I  
site and the oligo-dT track which allows identification of  
the library of origin of a clone within the mixture."

BASE COUNT 106 a 105 c 101 g 144 t  
ORIGIN

Query Match 9.5%; Score 95; DB 10; Length 456;  
Best Local Similarity 54.8%; Pred. No. 4e-17;  
Matches 210; Conservative 0; Mismatches 170; Indels 3; Gaps 1;

618 gtacaaaagaactctgtttagaatcagatcaccttccacatcgctgaagcag 677  
439 GTATAAGAGAGCCCTTATCTGATACACAGATGACCTTTAGAGAGACTATGACAA 380  
678 gccacccctgtagtgctgtagtccatccatccagctgaggttatgatacagtgct 737  
379 ACCAGCCTGGATGGGCTTTTAAAGCCCTGTCATCAAAAGCTTGAAGTGAACAAATTTC 320  
738 catcacactaaagaatgctcccatcctgtagctcagctcattgctgtagtacctta 797  
319 TGTTCACGTAAAGACTTTGCTTAGGCCCTTACACTTTTCTCATGAGGCTTACTTA 260  
798 ctggaagctctctgaggaagctggaatagatagacacagctcaaaaggagaagaaga 857  
259 CACCAAGGCGAAGAGAGGAGGAGCCATACCTTGACCAACACACTGATTTTCAAGAGCGCA 200  
858 tgaataagctctccctgggaggaagcattacatattgctggaagctccggaagaagatg 917  
199 TGACAAAGCTTTTCTGAGACACAGATTTTGAAGCTGCTGGTGGC---AATGAGCCAAAG 143  
918 tccaatgacctctgagccacactgtgcttactaccatcatattcttcacatgagccctgct 977  
142 TCTTGCGAGGAGGAGACAGCAATTTGTGACCAAGATTACACACTCTCATGATGCTGCC 83  
978 aaaagcttgaattacagcctca 1000  
82 AAAAGATTATGATGATGAGACTCA 60

RESULT 3  
BF150467 669 bp mRNA EST 29-DEC-2000  
LOCUS  
DEFINITION  
uy86b07.y1 NCI-GAP\_Mam5 Mus musculus cDNA clone IMAGE:3666421 5'  
similar to SW:CEMU\_MOUSE 061147 CERULOPLASMIN PRECURSOR, mRNA  
sequence.  
ACCESSION  
BF150467















DB 559 CAAAGATATTGCCTCAGACACTCA 582

||||| | ||||| fff  
RESULT 13  
BG6862595  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

          BG6862595                 632 bp          mRNA                 EST  
          G027959578F1 NIH\_CGAP\_Mam4 Mus musculus cDNA clone IMAGE:4917258 5'  
          BG6862595  
          BG6862595.1 GI:14213133  
          EST.  
          house mouse.  
          Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 632)  
NIH-http://nigc.mcl.nih.gov/.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
unpublished (1999)  
Contact: Robert Strausberg, Ph.D.

Plate: L1AMI0828 row: k column: 19  
High quality sequence stop: 631.  
Location/Qualifiers  
1  
538

```

/organism="Mus musculus"
/db_xref="taxon:10090"
/clone IMAGE:4917258
/clone_lib="NIH_CGAP_Mam4"
/tissue_type="tumor, gross tissue"
/lab_host="DRI0B"
/ncbi_organ="mammary; Vector: PCMV-SpOrf6; Site:1; Not1; Site:2; Salt: Cloned unidirectionally. Prime: Oligo dr. Average insert 2.5 kb. Library constructed by Life Technologies, catalog # 12018-016. Investigators providing samples: Lohrer, Hennighausen/Priscilla Furth, NIH Reference for transgenic model: Li et al., Cell Growth and Differentiation 7, 3-11 (1996). Note: this is a NCI_CGAP Library."

```

Query Match	9.1%	Score 90.8	DB 11	Length 632
Best Local Similarity	54.2%	Pred. No. 7	9e-16	
Matches 207	Conservative	0	Mismatches 172	Indels 3
			Gaps	1
QY	619	tacaaagaactctgtttagataatcacagatcaacttttcaacatcgctaaagc	678	
Db	210	TTATTAAGAGGCCCTTTATTTTGTAGTACACAGATGGCACCTTTAGTCACACTATAGACANA	269	
QY	679	caacctgatgtgctgtcgttagtctctaccacacagcgctgaagttatgatcaagtgtc	738	
Db	270	CCAGCGTGGCTATGGGTTTTTAGGCCCTTGTCATCAACGAAGGTGAACATATAAGTTAT	329	
QY	739	attacaacttaagaacatggtcttccatccctgtcaagtcctgatgctgtgtgttatctac	798	
Db	330	GTTCACTTAAACAACTTCCCTCTGTGGATCTTACACTTTTTCATGCACATGGGTACGTAC	389	
QY	799	tggaaagcttcctcgagggagctgaatatgatgatctcgaccacagtcgaaggagaagaagt	858	
Db	390	ACCAAGGAGTATGAGGAGGCCGTCTACCGCTAGACAAACCACTGATTTTTCACGGGCTGAT	449	
QY	859	gataaagcttccctcggttggaagccatacatatgtctgtgaagtccttgaagaagaatgt	918	

Db	450	GACAAAGTGCTCCCGGACACAGATGTGATGTGGCGCATGCC	---	AATGACCAAGT	506
Qy	919	ccaaaggcccttcagaccacvtgacctaacctacatactcttcata	tgtygaactgta	978	
Db	507	CCTGGAGGAGGAGACAGCAATTTGTGTGACGAGTATTACACTCCCACTCCAGTGTGATGCTCCA		566	
Qy	979	aaagactgaaatlcagagcctca	1000		
Db	567	AAAGTATTTCATCAGCACTCA	588		

RESULT	14
AV653260	AV653260 634 bp mRNA EST 07-SEP-2000
LOCUS	AV653260 GLC Homo sapiens CDNA clone GLCDA09 3', mRNA sequence
DEFINITION	AV653260
ACCESSION	AV653260.1 GI:9874274
VERSION	EST.
KEYWORDS	human.
SOURCE	
ORGANISM	

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Burakayota; Metazoa; Chordata; Crinacea; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Carnivora; Homnidae; Homo. 1 (bases 1 to 634)	Qian, B., Wu, T., Huang, Q., Kang, B., Gao, X., Xu, Z., Xiao, H., Xu, X., Li, N., Peng, Y., Liu, F., Qu, J., Song, H., Cheng, Z., Qu, J., Zeng, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G., Yang, Y., Gu, Y., Chen, Z., and Han, Z.	Homo sapiens cDNA clone	Unpublished (2000)	Contact: Zenguan Han

CHINESE NATIONAL HUMAN GENOME CENTER at Shanghai  
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai  
201203, P. R. China  
Tel: 86-21-50801919(ex. 45)  
Fax: 86-21-50801922  
Email: hanzq@chgc.sh.cn  
This clone is available at CHGC in Shanghai.  
Location/Qualifiers  
1..634

BASE COUNT	BRIGIN
213	a 119 c 128 g 172 t 2 others

	Query Match	Score	DB ID	Length
	Best Local Similarity	9.0%;		634;
	Matches	200; Conservative	52.18; Pred. No. 1.8e-15;	Mismatches 184; Indels 0; Gaps 0;
QY	617	lgtcaaaaagacactcgttgttagaattlaccagatcacaccttccaacatcgtaagccaa	676	
Db	203	TATATAAGAGGCCCTTTATCTTCAGTACACAGATGAACCTTTAGGCAACATATAGSAA	262	
QY	677	gcccacctgagatcggtctgcgtagatcctaccatccagctgaggtttagatacagttg	736	
Db	263	AACCGGTGTGGCTTTTAGGCCCTATTATCAAAAGCTGAACAGATPAAGATT	322	
QY	737	tcaataacttlaagaacatgcttcccatccctgcgtacgtcttcaatcgctgtgltacct	796	
Db	323	ATGTAACTTAAAAAACCTTGGCTTAGGCCCTACACCTTTCATTCACATGAAATAACTT	382	
QY	797	actggaagaacttcctggagagctgtaatatgtagtcagaccagtcgaaggagaagaag	856	
Db	383	ACTATAAGGACATGAGGGGGCCATCTACCTGATTAACACCCACAGATTTTCAAGAGCGAG	442	

Accession	Sequence	Position
QY	atgataagctctccctcgttggaaagccagatcacatattcgtgaagtcctgaagaagatg	916
Db		
443	ATGACCAAGATATTTTCAGGAGAGCAGTATCATATCATATGTTGCTTCCCACTAAGAACACA	502
QY	gtccaatgacctcagaccacactgtgcctacacacacatcatctctctatgttgaccttg	976
Db		
503	GTCCTGGGGAGAGAGATGGCAATTCTGTGACTGTAGGATTTAACCTTCCACATCATGATGCTC	562
QY	taaaagacttgaattcgaagcctca	1000
Db		
563	CAAAAGATNTGGCTTCAGACTCA	586

RESULT 15	880 bp	mrna	EST	24-OCT-2000
BF144665				
LOCUS	BF144665			
DEFINITION	601791527F1 NCL_CGAP_Lu30 Mus musculus cdna clone IMAGE:4022022 5'			

ACCESSION	BF144665
VERSION	BF144665.1
	GI:10983705

SOURCE	ORGANISM
house mouse.	Mus musculus

REFERENCE	1 (bases 1 to 860)
AUTHORS	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1995)
COMMENT	Contact: Robert Strausberg, Ph.D. <a href="mailto:robertstra@mail.nih.gov">robertstra@mail.nih.gov</a>

Email: cygpus@remail.org  
 Tissue Procurement: Gilbert Smith, Ph.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: Imple Genomics, Inc.  
 Clone distribution: NCICGAP clone distribution  
 information can be found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LLAM5278 row: f column: 07  
 High quality sequence: 679.

FEATURES  
source

Page 10

**"Music Made Me"**

---

10

BASE COUNT	268 a	150 c	236 g	220 l
ORIGIN				

Query Match	8.7%;	Score 87.2;	DB 11;	Length 880;
Best Local Similarity	54.0%;	Pred. No. 1e-14;		
Matches 20;	Conservative	0;	Mismatches 168;	Indels 3;
				Gaps 1.

[illegible]

Qy	799	tggaaagcttcctgaggaagctcgataatgatacgaaacccagcaaaaggagaaagaagat	858
Db	393	ACCACAGAGATGAGGGAGCCGCTACCCCTGACACACCACTGATTTTTCACGGCGTGAT	452
Qy	859	gataaagcttcctccctgaggaagcaatataatgctctggcaggtccctgaagagatatgct	918
Db	453	GACAAAGTGTTCGCCGACAAACAGATGTGATATGTGCTGCATGCC---AATGAGCCAAGT	509
Qy	919	ccaaatggcctctgagcccaactgtgcttaacttaactcaatcttctcaatgctgagactgta	978
Db	510	CTGTGAGAGAGGAGACAGCAATTTGTGTGACCGAGATTTTACCACTTCCCATGTTGATGTCGA	565
Qy	979	aaagacttgaat	990
Db	570	AAAGATTTTGAT	581

Search completed: January 8, 2002, 15:12:55  
Job time: 4771 sec

1



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 8, 2002, 17:15:54 ; Search time 7258.17 Seconds  
(without alignments)  
2275.185 Million cell updates/sec

Title: US-09-740-211-13\_COPY\_6000\_7000

Perfect score: 1001  
Sequence: 1 ccgtgcacttaccgcctc.....ggctggcttaactatgcgcgc 1001

Scoring table: IDENTITY\_NUC  
Gapop 10.0', Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl: 1: gb\_da: 2: gb\_htg: 3: gb\_in: 4: gb\_om: 5: gb\_ov: 6: gb\_pat: 7: gb\_ph: 8: gb\_pl: 9: gb\_pr: 10: gb\_ro: 11: gb\_sts: 12: gb\_sy: 13: gb\_un: 14: gb\_vi: 15: em\_da: 16: em\_fun: 17: em\_hum: 18: em\_in: 19: em\_om: 20: em\_or: 21: em\_ov: 22: em\_pat: 23: em\_ph: 24: em\_pl: 25: em\_ro: 26: em\_sts: 27: em\_sy: 28: em\_un: 29: em\_vi: 30: em\_htgo\_hum: 31: em\_htgo\_inv: 32: em\_htgo\_rnd: 33: em\_htg\_hum: 34: em\_htg\_inv: 35: em\_htg\_rnd: 36: em\_htg\_other:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	1001	100.0	2297	6 AX027819	AX027819 Sequence
c 2	1001	100.0	2368	12 AF264618	AF264618 Cloning vec
c 3	1001	100.0	2417	12 CVPSP70	X65330 Cloning vec
c 4	1001	100.0	2419	12 CVPSP71	X65331 Cloning vec
c 5	1001	100.0	2450	12 CVU74374	U74374 Cloning vec
c 6	1001	100.0	2462	12 CVPSP72	X65332 Cloning vec
c 7	1001	100.0	2464	12 CVPSP73	X65333 Cloning vec
c 8	1001	100.0	2512	12 ASAJ2679	AJ002679 Synthetic
c 9	1001	100.0	2665	12 SYNPU8V	L08959 pUC8 cloning
c 10	1001	100.0	2665	12 SYNPU8V	L09128 pUC9 cloning
c 11	1001	100.0	2665	12 SYNPU8V	L08880 pUC7 cloning
c 12	1001	100.0	2674	12 SYNPU8V	L08958 pUC7 cloning
c 13	1001	100.0	2680	12 SYNPU8V	A02712 pUC13 DNA s
c 14	1001	100.0	2680	12 SYNPU8V	L09129 pUC12 clone1
c 15	1001	100.0	2686	12 SYNPU8V	L09130 Cloning vec
c 16	1001	100.0	2686	12 SYNPU8V	A02710
c 17	1001	100.0	2686	12 SYNPU8V	L08752 pUC18 clone1
c 18	1001	100.0	2695	12 SYNPU8V	M77789 Cloning vec
c 19	1001	100.0	2695	12 SYNPU8V	L08911 pUC119R clo
c 20	1001	100.0	2695	12 SYNPU8V	L09138 pUC8-1 expr
c 21	1001	100.0	2695	12 SYNPU8V	L09139 pUC8-2 expr
c 22	1001	100.0	2695	12 SYNPU8V	L09140 pUC9-1 expr
c 23	1001	100.0	2695	12 SYNPU8V	L09141 pUC9-2 expr
c 24	1001	100.0	2701	12 SYNPU8V	L08910 pUC119H clo
c 25	1001	100.0	2704	12 SYNPU8V	L08941 POM9 cloning
c 26	1001	100.0	2704	12 SYNPU8V	L09142 pUC830 clone
c 27	1001	100.0	2710	12 SYNPU8V	L09143 pUC931 clone
c 28	1001	100.0	2710	12 SYNPU8V	L08939 POM3 cloning
c 29	1001	100.0	2713	12 SYNPU8V	L08940 POM4 cloning
c 30	1001	100.0	2713	12 SYNPU8V	L08937 POM1 cloning
c 31	1001	100.0	2716	12 SYNPU8V	L08938 POM2 cloning
c 32	1001	100.0	2716	12 SYNPU8V	L08912 pUC20H clone
c 33	1001	100.0	2722	12 SYNPU8V	X65304 Cloning vec
c 34	1001	100.0	2722	12 SYNPU8V	X65305 Cloning vec
c 35	1001	100.0	2743	12 CVPSP72	U34922 Cloning vec
c 36	1001	100.0	2746	12 CVPSP72	X65302 Cloning vec
c 37	1001	100.0	2854	12 CVPSP72	L08870 pGEM1 clone1
c 38	1001	100.0	2865	12 CVPSP72	X65302 Cloning vec
c 39	1001	100.0	2867	12 CVPSP72	L08872 pGEM3 clone1
c 40	1001	100.0	2867	12 CVPSP72	X65301 Cloning vec
c 41	1001	100.0	2869	12 CVPSP72	L08871 pGEM2 clone1
c 42	1001	100.0	2869	12 CVPSP72	X65303 Cloning vec
c 43	1001	100.0	2871	12 CVPSP72	U25273 Expression
c 44	1001	100.0	2891	12 CVPSP72	
c 45	1001	100.0	2891	12 CVPSP72	

## ALIGNMENTS

RESULT 1  
LOCUS AX027819/c 2297 bp DNA PAT 16-SEP-2000  
DEFINITION Sequence 11 from Patent WO0039307.  
ACCESSION AX027819  
VERSION AX027819.1 GI:10188663

KEYWORDS  
SOURCE  
ORGANISM  
synthetic construct.  
artificial sequence.

REFERENCE  
1 (bases 1 to 2297)  
Bestetti, G., Cali, S., Orsini, G., Tonon, G., Zuffi, G. and Ghisotti, D.  
Recombinant bacterial strains for the production of natural  
nucleosides and modified analogues thereof

JOURNAL  
BESTETTI GIUSEPPINA (IT) ; CALI SIMONA (IT) ; ORSINI GABRIELE (IT) ;  
ORSONI GABRIANO (IT) ; TONON GIANCARLO (IT) ; ZUFFI GABRIELE (IT) ;  
GHISOTTI DANIELA (IT)  
PATENT: WO 0039307-A1 06-JUL-2000;  
GHSOTTI DANIELA (IT)  
Location/Qualifiers  
1..2297

FEATURES  
source

/organism="synthetic construct"  
 /db\_xref="taxon:32630"  
 /note="cloning vector derived from pUC18"  
 BASE COUNT 571 a 574 c 580 g 572 t  
 ORIGIN

Query Match 100.0%; Score 1001; DB 6; Length 2297;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-274;  
 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 cctgaacttaccgctccatccatccatctatctgttgcgggaagctagaagt 60
Db 1142 cctgaacttaccgctccatccatccatctatctgttgcgggaagctagaagt 1083
QY 61 agtcgcagtaataagtttgcgaacgtttgccaatgacatcggtgttca 120
Db 1082 agttcgccagttatattggttgcgaacgtttgccaatgacatcggtgttca 1023
QY 121 cgttcgtcgtttgtatggttcaatcagtcgggtcccaagatcaaggcagttaca 180
Db 1022 cgttcgtcgtttgtatggttcaatcagtcgggtcccaagatcaaggcagttaca 963
QY 181 tgaatccccatgtttgcaaaaagcggttaagctccctcgttcctccgacgtttcaga 240
Db 962 tgaatccccatgtttgcaaaaagcggttaagctccctcgttcctccgacgtttcaga 903
QY 241 agtaagttgcgcagcagttatcaatcagttatggttgcgaacgtttccttcttact 300
Db 902 agtaagttgcgcagcagttatcaatcagttatggttgcgaacgtttccttcttact 843
QY 301 gtcacgtccacgttaagatgctttcttgcgtggtgagtaacctcaacgaatccttca 360
Db 842 gtcacgtccacgttaagatgctttcttgcgtggtgagtaacctcaacgaatccttca 783
QY 361 gaatagttatgctgcgaacgagttgctcttgcgcgagcgaatcaatgataatccgag 420
Db 782 gaatagttatgctgcgaacgagttgctcttgcgcgagcgaatcaatgataatccgag 723
QY 421 ccaatagcgggaacttaagaatgctcattcattgaaaaagttcctgcgggcaaaactc 480
Db 722 ccaatagcgggaacttaagaatgctcattcattgaaaaagttcctgcgggcaaaactc 663
QY 481 tcaagatcttaccgctgttgcagatccagttcagatgaacacacgttgcacccaatga 540
Db 662 tcaagatcttaccgctgttgcagatccagttcagatgaacacacgttgcacccaatga 603
QY 541 tcttcagcatcttacttaccacagcgttctcgtggtgagcaaaaacaggaagcaaat 600
Db 602 tcttcagcatcttacttaccacagcgttctcgtggtgagcaaaaacaggaagcaaat 543
QY 601 gccgcaaaaaggaataagggcgacaggaatgttgaatactactctctcttct 660
Db 542 gccgcaaaaaggaataagggcgacaggaatgttgaatactactctctcttct 483
QY 661 caatattatgaagatctatcaggttattgttcctatgaagcgaatataatttgaatgt 720
Db 482 caatattatgaagatctatcaggttattgttcctatgaagcgaatataatttgaatgt 423
QY 721 atttgaataaataaataaaggttccgcgacattcccccgaataagttgcacgtgac 780
Db 422 atttgaataaataaataaaggttccgcgacattcccccgaataagttgcacgtgac 363
QY 781 gcttaagaacaatctatctatcagatctatcaataaataaagcgtatcagagagccc 840
Db 362 gcttaagaacaatctatctatcagatctatcaataaataaagcgtatcagagagccc 303
QY 841 ttctgcctcgcgcttctgcgtatgacgltgaaaacccctacacatcagctcccgag 900
Db 302 ttctgcctcgcgcttctgcgtatgacgltgaaaacccctacacatcagctcccgag 243
QY 901 acgtgcacagcttgcgtatgacgltgaaaacccctacacatcagctcccgag 960

```

Db 242 ACAGTACAGCTTGTCTGTAAGCGGATGCCGGAGACAGACACCCGTCAGGCGCGTCA 183  
 QY 961 gcgggtgttgcgggtgtcggggcgtgcttaactatgcgcg 1001  
 Db 182 GCGGCTGTGCGGGGTGTGCGGGCTGACCTTAACATATGCCGC 142

RESULT 2  
 AF264618/c  
 LOCUS 2368 bp DNA circular SYN 03-OCT-2000  
 DEFINITION Cloning vector pUC6S, complete sequence.  
 ACCESSION AF264618 M74308  
 VERSION AF264618.1 GI:10504974  
 KEYWORDS

SOURCE  
 ORGANISM  
 Cloning vector pUC6S.  
 Cloning vector pUC6S.  
 artificial sequence; vectors.  
 1 (bases 1 to 2368)  
 2 (bases 1 to 2368)

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 MEDLINE  
 PUBMED  
 1905257  
 2 (bases 1 to 2368)  
 Vieira, J. and Messing, J.  
 Direct Submission  
 Submitted (27-APR-1993) Waksman Institute, Rutgers State  
 University, 190 Frelinghuysen Road, Piscataway, NJ 08854, USA  
 3 (bases 1 to 2368)  
 Vieira, J. and Messing, J.  
 Direct Submission  
 Submitted (26-SEP-2000) Waksman Institute, Rutgers State  
 University, 190 Frelinghuysen Road, Piscataway, NJ 08854, USA  
 Sequence update by submitter  
 On Oct 3, 2000 this sequence version replaced gi:209216.  
 Location/Qualifiers  
 1..2368  
 1..2368

REMARK  
 COMMENT  
 FEATURES  
 source  
 misc\_feature  
 1..127  
 /db\_xref="taxon:137702"  
 /note="multiple cloning site"  
 513..1373  
 /gene="bla"  
 513..1373  
 /gene="bla"  
 /codon\_start=1  
 /transl\_table=11  
 /product="beta-lactamase"  
 /protein\_id="AAG18332.1"  
 /db\_xref="GI:10504975"  
 /translation="MSIOHRYVALIPFAFCIPVAHPETLVKVKDAEDQLGARVGY  
 IELDLSNGKILESFRERPEPMWFKVLLCAVLRIDAGDEQLGRIRHNSDLVE  
 DRKEPVELNEAIPNDERDITTPVAMATLKLTLTGELTLASROQLIDWMEADKRVGLP  
 LRSALPAGWFIADRSAGRGSGRGIILALGPDGKRSRIVIVYITTSQATIMDRNRQIA  
 ELIGASLIKHW"

BASE COUNT 583 a 589 c 600 g 596 t  
 ORIGIN

Query Match 100.0%; Score 1001; DB 12; Length 2368;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-274;  
 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 cctgaacttaccgctccatccatccatctatctgttgcgggaagctagaagt 60
Db 1159 cctgaacttaccgctccatccatccatctatctgttgcgggaagctagaagt 1100
QY 61 agtcgcagtaataagtttgcgaacgtttgccaatgacatcggtgttca 120
Db 1099 agttcgccagttatattggttgcgaacgtttgccaatgacatcggtgttca 1040

```

QY 121 cgcctgctgttgatgagcttcattccagctccggttcccaagatcaaggcgattaca 180  
 Db 1039 CGCTCGTCTGTTGGTATGCTTCAAGCTCCGAGTCCCAAGATCAAGCGAGTTACA 980  
 QY 181 tgatcccccatgttgcagaaaggggttaagctcttccttcggtccggtcgttgcaga 240  
 Db 979 TGATCCCGCATGTTGTGCAAAAAAGGGGTGACCTCTTGCGCTCCATCTGCTTCTCGA 920  
 QY 241 agtaagttggcgcagtgatatactacatgattatgcagcactgacatcattcttact 300  
 Db 919 AGTAAGTTGGCGCCAGTGTATACATGATGATGATGATGATGATGATGATGATGAT 860  
 QY 301 gtaagcattcgttaagatgcttcttcgtgactggtgagatcctaaccagatcctcga 360  
 Db 859 GTCATGCCATCCGTAAGATGCTTTCTGTGACGTGTGATGATGATGATGATGATGAT 800  
 QY 361 gaataagttatggcgagcagatgctcttcgcccgcgtcaatacgggataataccgcg 420  
 Db 799 GAATAGTGTATGCGCGCATCCGAGTGTCTTGGCCGCGCATATACGGGATATACCGCG 740  
 QY 421 ccacatagcagaactttaaagttcattcattgggaaaaagttcttcggtggcgaaatc 480  
 Db 739 CCACATACAGAACTTTAAAGTCTCATGATGATGATGATGATGATGATGATGATGAT 680  
 QY 481 tcaagatcttaacgctggttggatgagatcagttcagttgaacccactggtgacccaactga 540  
 Db 679 TCAGAGATCTTACCGCTGTGTGATGATGATGATGATGATGATGATGATGATGAT 620  
 QY 541 tcttcagcattcttacttcttaccagcgttcttcgtgagcagaaagaaagaaagc 600  
 Db 619 TCTTACGATCTTACTTACTTACCGAGCTTCTGCTGTGAGCAAAAGAGAGCAAAAT 560  
 QY 601 ggcgcaaaaaggaataaggcgacacaggaatggtgaatcattcattccttctt 660  
 Db 559 GCCGCAAAAAGGAAATAGGCGACACGGAATGTTGAATGTTGAATGTTGAATGTT 500  
 QY 661 caaatatttgaagcatttcaatgaggttatttgcattcattgagcagatacatttgaatgt 720  
 Db 499 CAATATTTTGAAGCATTTATACAGGTTATGTCATGATGATGATGATGATGATGATGAT 440  
 QY 721 attagaataaataaataaaggttgcgcacattcccgaaaggtgacacactgac 780  
 Db 439 ATTTAATAAATAAATAAATAAAGGTTGCGGACATTTCCCGAAAGTCCACACTGAC 380  
 QY 781 gttcaagaacacattatcattgacatcaactaataaataagcgtatcaagagccc 840  
 Db 379 GTCTAAGAAACCATTTATGATGATGATGATGATGATGATGATGATGATGAT 320  
 QY 841 tttcgtctcgcggttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgt 900  
 Db 319 TTTTCGTCGCGGCTTTCGCTGATGATGATGATGATGATGATGATGATGATGAT 260  
 QY 901 aaggtcacagctgctcgttaagcgaatgcggagagcagaacagccgttaaggcgctga 960  
 Db 259 ACGGTCAACAGTGTCTGTAAGCGAGTCCGGGAGCAGCAAGCCGTCAGGGCGGTCA 200  
 QY 961 gcgaggtgttggcggtgttcggggcggtgcttaactatcgcgc 1001  
 Db 199 GCGGGTGTGGCGGGGTGCGGGCTGAGCTTAATGATGCGGC 159  
 RESULT 3  
 CVPSP70 2417 bp DNA circular SYN 16-JUL-1999  
 LOCUS CVPSP70 2417 bp DNA circular SYN 16-JUL-1999  
 DEFINITION Cloning vector pSP70.  
 ACCESSION X65330  
 VERSION X65330.1 GI:58235  
 KEYWORDS beta-lactamase; bla gene; cloning vector; multiple cloning site;  
 SOURCE Cloning vector pSP70.  
 ORGANISM Cloning vector pSP70  
 REFERENCE 1 (bases 1 to 2417)

AUTHORS Technical Services.  
 TITLE Direct Submission  
 JOURNAL Submitted (23-MAR-1992) Technical Services, Promega Corporation,  
 2800 Woods Hollow Road, Madison, WI 53711-5399, USA  
 REMARK 2 (bases 1 to 2417)  
 REFERENCE 1  
 AUTHORS Technical Services.  
 TITLE Direct Submission  
 JOURNAL Submitted (28-MAY-1993) Technical Services, Promega Corporation,  
 2800 Woods Hollow Road, Madison, WI 53711-5399, USA  
 COMMENT See X65330-X65335 for related vector sequences  
 This vector can be obtained from Promega Corporation, Madison, WI.  
 Information: One of the following numbers for order or technical  
 support:  
 In Wisconsin 800-356-9526  
 Outside U.S. 608-274-4330  
 Location/Qualifiers  
 1..2417  
 /organism="Cloning vector pSP70"  
 /db\_xref="taxon:90135"  
 FEATURES  
 source  
 misc\_feature  
 /note="Sp6 transcription initiation site"  
 4..45  
 /note="multiple cloning sites"  
 misc\_feature  
 /note="T7 transcription initiation site"  
 56..73  
 /note="T7 promoter"  
 complement(1090..1950)  
 /gene="bla"  
 complement(1090..1950)  
 /gene="bla"  
 /codon\_start=1  
 /product="beta-lactamase"  
 /protein\_id="CA46430.1"  
 /protein="GI:58236"  
 /db\_xref="GI:58236"  
 /translation="MSTQHRVALIPPEAFCLPYFAHPTLVKVAEDDGLGARVY  
 IELDINSKILIESRPREPEPMSTFVLLCGAVLSRIAGCGDGLRIHYSNDLVE  
 YSPVTEKHLIDGNTVRELCSAITSNTANLILITIGPKELTFLINMGDHYVRL  
 DRWPELNEAIPNDERDTPMVAATLRLILGELTLASROOLDMEADYVACPL  
 LRSALPAGWFIADKSGAGERSRGIILALPDGKPSRIVIVITTSQATWDERNRQIA  
 ELGSLIRKHW"  
 2401..2417  
 /note="Sp6 promoter"  
 BASE COUNT 615 a 599 c 592 g 611 t  
 ORIGIN  
 Query Match 100.0%; Score 1001; DB 12; Length 2417;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-274;  
 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 cctgcacatttaccgctccatccagctcattatgttgcgggaaagctagagtaag 60  
 Db 1304 CCTGCACCTTATCCCGCTCCATCCAGTCAATTAATGTTGCGGGAAGCTGAGTAAGT 1363  
 QY 61 agtcgacagttatagtttgcgcaagtggtgcatgctgacagcagcgtggtgca 120  
 Db 1364 AGTTGCGCAGTTAATGATGTTGCGCAAGCTGTTGCCATTGCTACAGCATCGTGGGTCA 1423  
 QY 121 cgcctcgtctgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgt 180  
 Db 1424 CGCTCGTGTGTTGGTATGAGCTTCAATGAGTCCGCTCCCAAGATCAAGCGAGTTACA 1483  
 QY 181 tgatcccccatgttgcagaaaggggttaagctcttccttcggtccggtcgttgcaga 240  
 Db 1484 TGATCCCGCATGTTGTGCAAAAAAGGGGTGACCTCTTGCGCTCCATCTGCTTCTCGA 1543  
 QY 241 agtaagttggcgcagtgatatactacatgattatgcagcactgacatcattcttact 300  
 Db 1544 AGTAAGTTGGCGCGCATCCGAGTGTATGATGATGATGATGATGATGATGATGAT 1603

```

OY 301 gtcacatcgcgaagatgcttcttctgagctgagctacacacgaatctcttga 360
Db 1604 gtcacatcgcgaagatgcttcttctgagctgagctacacacgaatctcttga 1663
OY 361 gaataagtgatgagcgcgcgaagatgcttcttccgcgcgtcaatacaggaataacacg 420
Db 1664 gaataagtgatgagcgcgcgaagatgcttcttccgcgcgtcaatacaggaataacacg 1723
OY 421 ccacatagcagaacttaaaagtgctcatcatctggaagaaagcttccgcgcgcgaagaa 480
Db 1724 ccacatagcagaacttaaaagtgctcatcatctggaagaaagcttccgcgcgcgaagaa 1783
OY 481 tcaagagcttaccgcgcgttgaagatccagctgagtaacacacacgctgacacacacga 540
Db 1784 tcaagagcttaccgcgcgttgaagatccagctgagtaacacacacgctgacacacacga 1843
OY 541 tctcagacatcttactcttaccacagcttcttccgcgcgcgtgagcaaaaacaggaagca 600
Db 1844 tctcagacatcttactcttaccacagcttcttccgcgcgcgtgagcaaaaacaggaagca 1903
OY 601 gcccgaacaaaggaagaaagcgcgcgcgaagaaatgttgaatacataccttcttctt 660
Db 1904 gcccgaacaaaggaagaaagcgcgcgcgaagaaatgttgaatacataccttcttctt 1963
OY 661 caataatattgaagcattatcagcaggtatgtctcagagcgcgcgcgaagaaatgttga 720
Db 1964 caataatattgaagcattatcagcaggtatgtctcagagcgcgcgcgaagaaatgttga 2023
OY 721 atttagaataaataaacaataaggggtctcgcgcgcacatctccgcgaagaaatgttga 780
Db 2024 atttagaataaataaacaataaggggtctcgcgcgcacatctccgcgaagaaatgttga 2083
OY 781 gctaaagaacacattatcattatcagcaggtatgtctcagagcgcgcgcgaagaaatgttga 840
Db 2084 gctaaagaacacattatcattatcagcaggtatgtctcagagcgcgcgcgaagaaatgttga 2143
OY 841 ttctgctcgcgcgttctcgtgagcgcgcgcgaagaaatgttgaatacataccttcttctt 900
Db 2144 ttctgctcgcgcgttctcgtgagcgcgcgcgaagaaatgttgaatacataccttcttctt 9203
OY 901 acggtcacagcgttctcgtgagcgcgcgcgaagaaatgttgaatacataccttcttctt 960
Db 2204 acggtcacagcgttctcgtgagcgcgcgcgaagaaatgttgaatacataccttcttctt 2263
OY 961 gggggtgctgagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1001
Db 2264 gggggtgctgagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 2304

```

```

FEATURES
Source
misc_feature
misc_feature
promoter
misc_feature
gene
CDS
This vector can be obtained from Promega Corporation, Madison, WI.
Call one of the following numbers for order or technical
information:
Order or Technical 800-356-9526
In Wisconsin 800-356-9526
Outside U.S. 608-274-4330.
location/Qualifiers
1. 2419
/organism="Cloning vector pSP71"
/db_xref="taxon:90136"
1
/note="SP6 transcription initiation site"
6. 47
/note="multiple cloning sites"
53. 75
/note="T7 promoter"
58
/note="T7 transcription initiation site"
complement(1092..1952)
/gene="bla"
complement(1092..1952)
/gene="bla"
/codon_start=1
/transl_table=1
/product="beta-lactamase"
/protein_id="CA46431.1"
/db_xref="GI:58238"
/translation="MSIQHFRVALIIFPAFCLEPFAHPETLVKVKDAEDQLGARVGY
IELDLSGKILIEFRELPRMSTFEVLGAVLIRIDAGQOLGRIRHYSONDLVE
YSPVTEKHLDIGMTVRELCSAITSMSNTANLLITIGSRKELTRFLNNMDVTRL
DRMPETLENATPNDERDTPVAMATLRKLTLGELLTLASRQDLIDMEADKVGPL
ELGASLIKHHN"
2402..2419
/note="SP6 promoter"
BASE COUNT 615 a 601 c 592 g 611 t
ORIGIN
Query Match 100.0%; Score 1001; DB 12; Length 2419;
Best Local Similarity 100.0%; Pred. No. 1.2e-274;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 cctgcacattatccgcctccatccacgaatcttgaatgttgcgcgcgaagaaatgttga 60
Db 1306 cctgcacattatccgcctccatccacgaatcttgaatgttgcgcgcgaagaaatgttga 1365
OY 61 agtcgcgaagtaagattgagcgaagcgttgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 120
Db 1366 agtcgcgaagtaagattgagcgaagcgttgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1425
OY 121 cgtcgtcgttctgagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 180
Db 1426 cgtcgtcgttctgagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1485
OY 181 tgaatccccaatgttgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 240
Db 1486 tgaatccccaatgttgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1545
OY 241 agtaagtgccgcgaagtgatcactcagtgatgagcgaagcgcgcgcgcgcgcgcgcgcgc 300
Db 1546 agtaagtgccgcgaagtgatcactcagtgatgagcgaagcgcgcgcgcgcgcgcgcgcgc 1605
OY 301 gtcacatcgcgaagatgcttcttctgagctgagctacacacgaatctcttga 360
Db 1606 gtcacatcgcgaagatgcttcttctgagctgagctacacacgaatctcttga 1665
OY 361 gaataagtgatgagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 420
Db 1666 gaataagtgatgagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1725
OY 421 ccacatagcagaacttaaaagtgctcatctggaagaaagcttccgcgcgcgcgcgcgcgcgcgc 480

```

COMMENT  
See X65300-X65335 for related vector sequences

Db 1726 CCACATAGCAAGACTTTAAAGTGTCTCATCATGAAAAAGCTTTCTGGGGCGAAGAACTC 1785  
 QY 481 tcaagatcttaccgctgttgagatccagttcgatgtaaccactcgtgcacccaactga 540  
 Db 1786 TCAAGGATCTTACCGCTGTGAGATCCAGTTCGATGTAACCCACATCGTGACACCCAACTGA 1845  
 QY 541 tcttcagcatcttcttaccctaccagcgtttcttggttgagcaaaacaggaagcaaat 600  
 Db 1846 TCTTCAGCATCTTTACTTCTTACACACCGTTCCTGGGTGAGCAAAAACAGGAAGCAAAAT 1905  
 QY 601 gccgcaaaaaaggaataaaggcgacacaggaatgttgaatctcactcttctt 660  
 Db 1906 GCCGCAAAAAAGGAATAAGGCGACACGGAATGTGTAATCTCACTCTCTCTTTT 1965  
 QY 661 caatattatgaagcatltaacaggttatctgtctcaatgagggatataatttgaat 720  
 Db 1966 CAATATTATTGAACATTTATTCAGGGTTATTGTCATGACGGATACATATTGTAATGT 2025  
 QY 721 attagaaaaataaacaagaggggttcgcgcacatctcccgaaaaagtgccactgac 780  
 Db 2026 ATTTAGAAAAATTAACAATAGGGCTTCGCCGACATTTCCCGAAAGTGCCACTGAC 2085  
 QY 781 gtctagaagaacatattatcatatgacatgaacataaaatagcggtatcaagaagccc 840  
 Db 2086 GCTTAAGAAACCATATATATCATGACATTAACCTTAATAAGGCTATCACGAGGCC 2145  
 QY 841 ttctgctcgcgcttctggtgatgagtgagaaacctctgacacatgcagctcccgag 900  
 Db 2146 TTTCTCTTCGCGCGTTCGGTGATGACGGTAAACCTCTGACACATGACGCTCCGAG 2205  
 QY 901 acgggtcacagctctgtctgaagcgatgcgagacgaacacgcgtcaggcgagtc 960  
 Db 2206 ACGGTACACGCTGTCTGAAGCGGATCGCGGAGACAGAACCCGTCAGGCGCGTCA 2265  
 QY 961 gcgggtgttgccgggtgttcggggcggttaactatgcgc 1001  
 Db 2266 GCGGGTGTGGCGGGTGTGGGGCTGAGCTTAACCTATGCGGC 2306  
 RESULT 5  
 CVU74374 2450 bp DNA circular SYN 29-OCT-1996  
 LOCUS CVU74374 2450 bp DNA circular SYN 29-OCT-1996  
 DEFINITION Cloning vector pUF5, complete sequence.  
 ACCESSION U74374  
 VERSION 074374.1 GI:1649038  
 KEYWORDS  
 SOURCE Cloning Vector pUF5.  
 ORGANISM Cloning Vector pUF5.  
 ARTIFICIAL SEQUENCE: vectors.  
 REFERENCE 1 (bases 1 to 2450)  
 AUTHORS Fan, J.-B., Quackenbush, J. and Myers, R.M.  
 JOURNAL pUF5 Cloning Vector  
 REFERENCE 2 (bases 1 to 2450)  
 AUTHORS Fan, J.-B., Quackenbush, J. and Myers, R.M.  
 JOURNAL Direct Submission  
 TITLE Submitted (11-OCT-1996) Stanford University, Stanford Human Genome  
 COMMENT Center, 855 California Avenue, Palo Alto, CA 94304, USA  
 CONTACT: R. M. Myers  
 Stanford Human Genome Center  
 855 California Ave.  
 Palo Alto, CA 94304 USA.  
 Ampicillin Resistance Selectable Vector used in Transposon-Mediated  
 Sequencing.  
 FEATURES  
 source  
 1. .2450  
 /organism="Cloning Vector pUF5"  
 /db\_xref="taxon:33561"  
 misc-feature  
 1  
 /note="SP6 transcription initiation site"  
 4. .78  
 /note="Multiple Cloning Site  
 KhoI/PvuII/BstXI/XbaI/BamHI/EcoRI/BstXI/EcoRV/BglII"

promoter 86..106  
 /note="T7 promoter"  
 misc-feature 88  
 /note="T7 transcription initiation site"  
 CDS complement(1123..1983)  
 /codon\_start=1  
 /transl\_table=11  
 /product="beta-lactamase"  
 /protein\_id="AB17670.1"  
 /db\_xref="GI:1649039"  
 /translation="MSIQHFRVALIPEFAACLPVPAHPETLVKKADEQLGARVY  
 IELDLSNGKILLESFRRPEERFPMSTFKVLICGAVLSRIDAGQDQLGRIRYSDNLYE  
 VSPETKEKLLDGMIVRELCSAIIITMSDITANLLITLTGSKELFALDHMGDIYRL  
 DRIPELNLAIINDESDITMPVAAATYLRKLLTGLTEGLTIAERQOLIDMEADKVGFL  
 LNSALPAGWFIADKSGAGRSRGIIALGPDKPSRIIVLYITTSQATMDERNROLA  
 EIGASLIKHW"  
 2434..2450  
 /note="SP6 promoter"  
 BASE COUNT 624 a 609 c 602 g 615 t  
 ORIGIN  
 Query Match 100.0%; Score 1001; DB 12; Length 2450;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-274; Indels 0; Gaps 0;  
 Matches 1001; Conservative 0; Mismatches 0;  
 QY 1 cctggaacttaccgcgcctccatccagctcatatgaattgttcgcgggaagctagatgaat 60  
 Db 1337 CCGTGAACCTTTATCCGCCCTCCATCCAGTCTATTAATGTTGCGGGAAAGCTAGATTAAGT 1396  
 QY 61 agtcgcagtaataagttctgcgaagttgttgccatgcttaacagcagtcggtgtca 120  
 Db 1397 AGTTCGCAGTTAAATGATGTTGCGCAACGTTGTGCTGCTACAGGCAATCGTGGTCA 1456  
 QY 121 cgcctgcgtgttgatgagcttcatcagctcgcggttcccaagatcaagcgagttaca 180  
 Db 1457 CGCTCGCTGTTGATGATGCTTCATTCAGCTCCGGTCCCAACGATCAAGCGAGTTACA 1516  
 QY 181 tgatcccccatgtgtgtcaaaaaagcggttagctcttcgcttcgctccatcggtgtcaga 240  
 Db 1517 TGAATCCCATGATGTTGCGCAAAAAGCGGTGACTCTCTTCGCTCCGATCGCTGTCAGA 1576  
 QY 241 agtaagttggcgcaatgtgtatcaatcagttatgtatgtagcaagcagtcataatctctact 300  
 Db 1577 AGTAAGTTGGCGGCAATGTTATCATCATGATGATGATGATGATGATGATGATGATGAT 1636  
 QY 301 gtaatgccalcgtaagatgtcttctgtgactggttgagactcaaccaagtcattctga 360  
 Db 1637 GTCATGCCATCCGTAGATGCTTTCTGTGACTGGTGAATCTCAACCAAGTCAATTCGA 1696  
 QY 361 gaataagttatggcgcaacgagttgctcttcgcccgcgtcaatacgggataataccgcg 420  
 Db 1697 GAATAGTGTATGGCGGCAACGAGTTCCTGTCGCCGCGCTCAATACGGGATTAATACCGCG 1756  
 QY 421 ccacatagcagaacttaaaagtgtctcatcaatctggaaaaagttcttcggggcgcaaaact 480  
 Db 1757 CCACATAGCAGAACTTTAAAGTGTATCATGTAAGAAAGTTCCTGCGGGCAAAACTC 1816  
 QY 481 tcaagatcttaccgctgttgagatccagttcgatgtlaaaccactcgtgcaccaactga 540  
 Db 1817 TCAAGGATCTTACCGCTGTGATGATCCAGTTCGATGAATGAACCCATCGTCCACCAACTGA 1876  
 QY 541 tcttcagcatcttcttaccctaccagcgtttcttggttgagcaaaacaggaagcaaat 600  
 Db 1877 TCTTCAGCATCTTTACTTCTTACACACCGTTCCTGGGTGAGCAAAAACAGGAAGCAAAAT 1936  
 QY 601 gccgcaaaaaaggaataaaggcgacacaggaatgttgaatctcactcttctt 660  
 Db 1937 GCCGCAAAAAAGGAATAAGGCGACACGGAATGTGTAATCTCACTCTCTCTTTT 1996  
 QY 661 caatattatgaagcatltaacaggttatctgtctcaatgagggatataatttgaat 720  
 Db 1997 CAATATTATTGAACATTTATTCAGGGTTATTGTCATGACGGATACATATTGTAATGT 2056

Qy 721 attagaaataaacaataggggtccgcgcacattcccccgaagaagtgcacactgac 780  
 Db 2057 ATTGGAATAAATAAACAATAGGGGTTCCGGCACATTTCCCAAAAGTGCCACCTGAC 2116  
 Qy 781 gcttaagaacacattatcatgacattacataaataaggcgatcacagaagccc 840  
 Db 2117 GCTTAAGAACAACATTATATCATGACATTAAACCTATTAATAATAGGCGTATCACAGAGCCC 2176  
 Qy 841 ttcttcgcgcgcgttccttcgtatgacggtgaaacccctgcacacatgcagcccgag 900  
 Db 2177 TTCTGCTCCGCGGTTCCTTCGATGATGACGTAAGAAACCTCTGACACATGCAAGCTCCCGGAG 2236  
 Qy 901 accgttcacagcttccttcgaagcagatgcggagcagaaagcccgtaaggcgatca 960  
 Db 2237 ACGGTACAGCTTCTCTGTAAGCGATGCGGAGACGACAAACCGCTCAGGCGCGTCA 2296  
 Qy 961 gcgggttgcgggggtgcgggctgagcttaactatgcgc 1001  
 Db 2297 GCGGGTGTGGCGGTGTCGGGCTGCTTAACCTATGCGGC 2337

RESULT 6  
 CVSP72  
 LOCUS CVPSP72 2462 bp DNA circular SYN 25-JAN-2000  
 DEFINITION Cloning vector psp72.  
 VERSION X65332  
 KEYWORDS beta-lactamase; bla gene; cloning vector; multiple cloning site;  
 promoter.

SOURCE Cloning vector psp72.  
 ORGANISM Cloning vector psp72.  
 REFERENCE 1 (bases 1 to 2462).  
 AUTHORS Technical Services.  
 TITLE Direct Submission

REMARK Submitted (23-MAR-1992) Technical Services, Promega Corporation,  
 2800 Woods Hollow Road, Madison, WI 53711-5399, USA

REMARK 2 (bases 1 to 2462)  
 AUTHORS Technical Services.  
 TITLE Direct Submission

REMARK Submitted (28-MAY-1993) Technical Services, Promega Corporation,  
 2800 Woods Hollow Road, Madison, WI 53711-5399, USA

REMARK 3 (bases 1 to 2462)  
 AUTHORS Technical Services.  
 TITLE Direct Submission

REMARK Submitted (12-JAN-2000) Technical Services, Promega Corporation,  
 2800 Woods Hollow Road, Madison, WI 53711-5399, USA  
 On Jan 26, 2000 this sequence version replaced g1:58239.  
 See X65300-X65335 for related vector sequences  
 Call one of the following numbers for order or technical  
 information:

Order or Technical 800-356-9526  
 In Wisconsin 800-356-9526  
 Outside U.S. 608-274-4330.

FEATURES  
 source Location/Qualifiers

1. 2462  
 /organism="Cloning vector psp72"  
 /db\_xref="taxon:90137"  
 misc\_feature 4..90  
 /note="SP6 transcription initiation site"  
 misc\_feature 99..118  
 /note="multiple cloning sites"  
 promoter 101  
 /note="T7 promoter"  
 misc\_feature 101  
 /note="T7 transcription initiation site"  
 gene complement(1135..1995)  
 /gene="bla"  
 complement(1135..1995)  
 CDS

/gene="bla"  
 /codon\_start=1  
 /transl\_table=11  
 /product="Beta-lactamase"  
 /protein\_id="CA46432.1"  
 /db\_xref="GI:58240"  
 /translation="MSIQHRYALIPFAAFLPVFAHPETLVKVKDAEDLAGRVY  
 IEIDLSGKILDESFEERFPKMSTFQVLLCGAVLSRIDAGEOLGRIRIHYSONDLYE  
 YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLTTGGKRELTAAFLHNGDVYTRL  
 DRWPELNAIPMDERDTTPVMAATTLTKLITGLTLASROOLIDMEADKVGPI  
 LRSALPWGFIADKSGAGERGSRGIIALGPGRFSRIVITITTSQATMDERNROLA  
 EIGASLIKRW"  
 John(246..2462,1..3)  
 /note="SP6 promoter."  
 BASE COUNT 624 a 614 c 605 g 619 t  
 ORIGIN

Query Match 100.0%; Score 1001; DB 12; Length 2462;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-274;  
 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ccgcacatttaccgctccatccagtcattatctgttcggggaagcctaagtaagt 60  
 Db 1349 CCGCAACTTTATCCCGCTCCATCCAGTCTATTGTTCCGGGAAGCTAGAGTAAGT 1408  
 Qy 61 agtgcacagtaataagtttgcaacgcttgcacatgcbaagcagtcgtgtca 120  
 Db 1409 AGTTCCCGAGTTAAATAGTTTGGCAACGTTGTGCAATTCCTACAGGCATGCTGTCTCA 1468  
 Qy 121 cgcctgcgttgcgtatgcgttcacatccagtcggttcgaagcagtcagttaca 180  
 Db 1469 CGCTGCTGCTTGTGATGAGCTTCATTCAGCTCCGCTCCCAACATCAAGCGAGTTACA 1528  
 Qy 181 tgatcccatctgttgcaaaaaaagcgttgccttcctgcgttcctccagtcgtgtca 240  
 Db 1529 TCATCCCCCATGTTGTGAAAAAGCGGTTAGCTTCCTCGCTCCGATCGTTGTACGA 1588  
 Qy 241 agtaagttgcgcagtglttcaactaagttatgtaagcagcactgataatcttact 300  
 Db 1589 AGTAAGTTGCGCGAGTGTATCATCTAGTTATGACACACTGCAATATCTCTTACT 1648  
 Qy 301 gtcacgcatccgtaagatgcttcttcgtgacgttgatgactcaacaaagtcactga 360  
 Db 1649 GTCATGCCATCCGTAAGATGCTTTCTGTGACTGTGAGTACCAACCAAGTCATTCTGA 1708  
 Qy 361 gaatagtgatgcgcagcagttgtcttcgcccggcggtcaataagcgaataacacgcg 420  
 Db 1709 GAATAGTGTATGCGCGGACCGAGAGTGTCTTCCCGCGCAATAGGAGTAAATACCGCG 1768  
 Qy 421 ccaatagcagaacttaaaagtctcatcatatggaacagtccttcggggcgaaacac 480  
 Db 1769 CCACATAGCGAAGACTTTAAAGTCTCATCATGTGAAGAGTCTTCCGGCGGCAAAACTC 1828  
 Qy 481 tcaaggtcttaccgctgttgatgatacagttgatacccaactcgtgcaccaactga 540  
 Db 1829 TCAAGGATCTTACCGCTGTGATGATCCAGTTCAGTAAACCCACATGAGGATTAATACCGCG 1888  
 Qy 541 tcttcagcactcttcaactcaacagcgttcttggttggaacaaacaggaagcgaat 600  
 Db 1889 TCTTCAGCATCTTTACTTTCACACACGTTCTGGGTGAGCAAAACAGGAAGCAAAAT 1948  
 Qy 601 gccgcaaaaaggaataaggcgcagacaggaatgtgaaatacctaactcttcctt 660  
 Db 1949 GCGGCAAAAAGGAATAGGGCGACAGCAAAATTTAAATACATCTCTTCTTTT 2008  
 Qy 661 caataltatgaagcatatataagggtatgtctcagcagcagatcatatgtgaatgt 720  
 Db 2009 CAATATTTTGAAGCATTTATCAGGTTATTTCTCATGACGAGATCAATTTGAATGT 2068  
 Qy 721 atttagaaaaataaacaataggggttcgcgcacattcccccgaagaagtgcacactgac 780  
 Db 2069 ATTGGAATAAATAAACAATAGGGGTTCCGGCACATTTCCCAAAAGTGCCACCTGAC 2128

QY 781 gctcaagaacacattatctatgacattacctaataaataagcgtatcacaagagccc 840  
 Db 2129 GTCATGAACAAACATTTATTCATGACATTAACCTTAATAAATAGCGCATCACAGAGCCC 2188  
 QY 841 ttctgctcgcggttcggtgtagacaggtgaacacctctgacacatgcagctccgagag 900  
 Db 2189 TTTCGTTCCGCGCTTTCGGTGTGATGACGGTGAACAACTCTGACATCACCCTCCCGAG 2248  
 QY 901 acggtcacagcttgctgtagacaggtgacaggtgacagagacagagccgctcagagcgctca 960  
 Db 2249 ACGGTCACAGCTGCTGTGTAAGCGGATGCCGAGACAGACAGCCCGCTCAGGCGCCGTCA 2308  
 QY 961 gcgaggtttgacgaggtgacgaggtgacgaggtgacgaggtgacgaggtgacgaggtgac 1001  
 Db 2309 GCGGCTGTTGGCGGCTGCGGCGCTGCTTAACCTAATGCGGC 2349  
 RESULT 7  
 LOCUS CVPSP73 2464 bp DNA circular SYN 25-JAN-2000  
 DEFINITION Cloning vector pSP73.  
 ACCESSION X65333  
 VERSION X65333.2 GI:6759495  
 KEYWORDS beta-lactamase; bla gene; cloning vector; multiple cloning site;  
 promoter.  
 SOURCE Cloning vector pSP73.  
 ORGANISM Cloning vector pSP73  
 REFERENCE 1 (bases 1 to 2464).  
 AUTHORS Technical Services.  
 TITLE Direct Submission.  
 JOURNAL Submitted (23-MAR-1992) Technical Services, Promega Corporation,  
 2800 Woods Hollow Road, Madison, WI 53711-5399, USA  
 REMARK 2 (bases 1 to 2464)  
 AUTHORS Technical Services.  
 JOURNAL Submitted (28-MAY-1993) Technical Services, Promega Corporation,  
 2800 Woods Hollow Road, Madison, WI 53711-5399, USA  
 REMARK 3 (bases 1 to 2464)  
 REFERENCE Technical Services.  
 JOURNAL Submitted (13-JAN-2000) Technical Services, Promega Corporation,  
 2800 Woods Hollow Road, Madison, WI 53711-5399, USA  
 COMMENT See X65300-X65335 for related vector sequences  
 This vector can be obtained from Promega Corporation, Madison, WI.  
 Call one of the following numbers for order or technical  
 information:  
 Order or Technical 800-356-9526  
 In Wisconsin 800-356-9526  
 Outside U.S. 608-274-4330.  
 FEATURES  
 source  
 1. 2464  
 /organism="Cloning vector pSP73"  
 /db\_xref="taxon:90138"  
 misc\_feature  
 1  
 /note="SP6 transcription initiation site"  
 6...92  
 /note="multiple cloning sites"  
 101...120  
 /note="T7 promoter"  
 103  
 /note="T7 transcription initiation site"  
 complement(1137...1997)  
 /gene="bla"  
 complement(1137...1997)  
 /gene="bla"  
 /cddon\_start=1  
 /transl\_table=11  
 /product="beta-lactamase"  
 misc\_feature  
 gene  
 CDS

/protein\_id="CAA6433.1"  
 /db\_xref="GI:36242"  
 /translation="MSIQHFRVALIPFAACILPFAHPETLVKVKDAEDQLGARVGY  
 IELDLSGKILDSFRPERFPMSTFKVLDCGAYSLRDNGEDELGRHYHSONDLVE  
 YSPVTEKHLLDGMETVRELCSAAITMSDNTANALFLITIGSKREKTFALHNKGHIVPL  
 DRPEPELNEAIPNDERDTTMVPMATMTLRKLRLGLLTLASRQDIIMHEADRYAGPL  
 LRSALPAGMTADIKSGAGERSGLIIAALGPDGKPSRLIVITVTGSSQATMDERRRQLA  
 EIGASLIKRW"  
 join(2448...2464,1...3)  
 promoter  
 /note="SP6 promoter"  
 BASE COUNT 624 a 615 c 606 g 619 t  
 ORIGIN

Query Match 100.0%; Score 1001; DB 12; Length 2464;

Best Local Similarity 100.0%; Pred. No. 1.2e-274; Mismatches 0; Gaps 0;

Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cctgcaacttaccgctccatccagctcatatattgttgcgaggaagctagaagtaagt 60  
 Db 1351 CCTGCAACTTTATCCCGCTCCATCCAGTCTATTATTGTTCCGGGAAGCTAGAGTAAGT 1410  
 QY 61 agtgcgcagtaataagttgagcaaacgtgttgcattgtcctatgacaggaatcgtgtgca 120  
 Db 1411 AGTTCGCCAGTTAATAGTTTGGCAACGTTTGGCATTGCTACAGGCATGCTGTGTCA 1470  
 QY 121 cgtctgcgttctgtagtgccttcatcagctccggttcccaagcaagcaagcgagttaca 180  
 Db 1471 CGCTTCGTGTTGGTATAGCTTCATTCACCTCCGCTCCCAACGATCAAGCGAGTTTACA 1530  
 QY 181 tgaatcccatgttgcgaataaagcggttagctcccttcgtccctccagatcgtgtcaga 240  
 Db 1531 TGATCCCCCATAGTTGTCAAAAAGCGGTTAGCTCTTGGTCCATCCATGCTGTGTCA 1590  
 QY 241 agtaagttgagcgaggttatacactaactaagttgtagcagcagctgataatcttact 300  
 Db 1591 ACTAAGTTGGCGCAGTGTATACATCATGTTATGACGACACTGCATATTCTTACT 1650  
 QY 301 gtcattgcacatcgttaagaatgcttctgtgacgaggtgagtaactcaaccaagtcattcga 360  
 Db 1651 GTCATGCCATCCGTAAAGATGCTTTCTGTGACGTGATGATCAACCAAGTCAATCTGA 1710  
 QY 361 gaatagttatgacgagcagcagatgctcttgcgagcgctcaatacggatataacgcg 420  
 Db 1711 GAATAGTGTATCGCGCAGCAGTGTGCTTCCCGCGTCAATACGGATTAATACCGG 1770  
 QY 421 ccaatagcagaacttaaaagtgtcaactaactgtaaaacgcttctcgagggaacac 480  
 Db 1771 CCACATAGCAGAACTTAAGATGCTCATCTAFTGGAACAGTCTTCCGGCGCAAAATC 1830  
 QY 481 tcaagatcttaccgctgttgagatcagctcgatgtaacccactgtgacccaactga 540  
 Db 1831 TCAAGATCTTACCGCTGTGATCCAGTTGATGATCAACCACTCGTGCACCACTGA 1890  
 QY 541 tcttcagcatcttcttacccttccacagcgttctcgggttagcgaataaaggaagcaaat 600  
 Db 1891 TCTTCAGCATCTTTTACTTTTCACACAGCTTCTCGGTGAGCAAAACGGAAGCAAAAT 1950  
 QY 601 gccgcataaaggaataaagggcgacagcgaagaaatgttgaataactaactcttctt 660  
 Db 1951 GCCGCAAAAAGGAATTAAGGCGACACGAAATGTTGATCTCATCTCTCTCTTTT 2010  
 QY 661 caatatatgaagcatttaccaggggttctgctcatagaggaataatattgaagt 720  
 Db 2011 CAATATTATTAAGCATTTTATCAGAGGTATTGCTCATGAGCGGATACATATTGAATGT 2070  
 QY 721 atttagaataaataaataaaggggttcgagcgaacttcccgaaaggtccactgac 780  
 Db 2071 ATTTAGAAATAAATAAATAGGGGTTCCGCGACATTTCCCGAATAAGTGCACCTGAC 2130  
 QY 781 gctagaagaacattatattatgacatlaaactataaataagcgctatcagagccc 840  
 Db 2131 GCTTAAGAAACCATTTATTCATGACATTAACTATAAATAGCGGTATCACAGAGCCC 2190

[illegible]



# AUTHORS

## TITLE

### JOURNAL

#### COMMENT

Gilbert, W.  
Obtained from Vecbase 3.0  
Unpublished (1991)  
These data and their annotation were supplied to GenBank by W111  
Gilbert under the auspices of the GenBank Curatorial Program. PUC8 -  
Cloning vector

#### #TYPE DNA CIRCULAR

ENTRY PUC8 - Cloning vector  
TITLE PUC8  
DATE 17-SEP-1986  
#sequence 16-DEC-1986  
ACCESSION V00021  
SOURCE artificial

#### REFERENCE

#number 1  
#authors Vieira J., Messing J.  
#journal Gene (1982) 19: 259-268  
#comment see also 'note added in proof'  
REFERENCE  
#number 2  
#authors Polkels P.H., Enger-Valk B.E., Brammar W.J.  
#book Cloning Vectors, Elsevier 1985 and supplements  
#comment Vector I-A-14-20

#### COMMENT

Assembled from pUC18 and M13mp8 by F. Pfeiffer, MPI, Martinsried  
Revised 16-DEC-1986 by F. Pfeiffer:  
2287/8 'AT' to 'TA' to match revised sequence of PBR322  
The strand shown contains the beta-galactosidase mRNA sequence  
including the multiple cloning site of M13mp8.

#### KEYWORDS

##### CROSSREFERENCE

#complement  
Vecbase(3):pUC8c  
#parent  
Vecbase(3):pUC7, Vecbase(3):M13mp8, Vecsource(3):bGal18  
#brother  
Vecbase(3):pUC9  
#ofspiring  
Vecbase(3):pUC12, Vecbase(3):PEMBL8m, Vecbase(3):PEMBL8p,  
Vecbase(3):pUC830, Vecbase(3):PM2, Vecbase(3):PM4,  
Vecbase(3):pCM8  
Vecbase(3):pUC81, Vecbase(3):pUC82, Vecbase(3):PIC7 PARENT  
Features of pUC8 (2665 bp)  
residue source  
1- 426 6002-6427 M13mp8  
1- 229 1069-1297 lac-Operon  
230- 265 1- 36 polylinker of M13mp8  
268- 426 M303-1461 lac-Operon  
427- 526 2351-2252 (c) PBR322  
527- 663 2210-2074 (c) PBR322  
664-2665 4355-2354 (c) PBR322  
Conflict (cfl) and Mutations (mut):  
pUC8 source  
mut 1107 T C 3912 (c) PBR322  
mut 1408 A G 3611 (c) PBR322

FEATURE  
931-1719 1-789 Ap-R, b-lactamase  
POLYLINKER EcoRI-SmaI-BamHI-SalI-PstI-HindIII  
SELECTION

#resistance Ap  
#indicator beta-galactosidase  
#puc8 #length 2665 #checksum 2136.  
SUMMARY Location/Qualifiers

1..2665  
/organism="synthetic construct"  
/db\_xref="taxon:32630"

#### BASE COUNT

#### ORIGIN

Query Match 100.0% Score 1001: DB 12: Length 2665:  
Best Local Similarity 100.0% Pred. No. 1.2e-274:  
Matches 1001: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

```

OY      1 cctgaacttattccgcctccatccagcttaattatgttgcgggaagcagaagtaag 60
DB      1510 CCGCAACTTTATCCGCCCTCCATCCAGTCTATTATTTCGCCGGGAAGCTAGAGT 1451
OY      61 agtgcagcagtaaatagtttgcgaacgtttgtccattgtcttaaggaatcagtggtca 120
DB      1450 AGTTCGCCAGTTAATAGTTTGGCAACGTTGTCATTGCTACAGGAGCTGTGTCTCA 1391
OY      121 cgtctcgtcttggtataggtcctcaatcagctccggttcccaacgataagcgagttaca 180
DB      1390 CGCTCGCTTGTGTGTATGGCTTCATTCACCTCCGGTCCCAACGATCAAGCGAGTTTACA 1331
OY      181 tgaatccccatgtgtgtgcacaaaagcggttagctccttcgtccctcagatgtgtcaga 240
DB      1330 TGAATCCCCATGTGTGTGCACAAAAGCGGTTAGCTCTTCGTCCATCGATGTTGTCTAGA 1271
OY      241 agtaagttgcgcagcagttgtatcactcaatgtgtatagcgaacgtcgaatattcttact 300
DB      1270 AGTAAGTTGCGCCGACGTTATCAGTCATGATGGAAGCACTGCATATATTCCTTACT 1211
OY      301 gtaatgcacatcgaatgaagcttctctgtgactgtgtgagtaactcaaccaagctattctga 360
DB      1210 GTCAATGCCATCCGTAAGATGCTTTCTGTGACTGTGTGACTACTCAACCAAGCATTTCTGA 1151
OY      361 gaatagtgatgcgcgacccagctgtccttgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 420
DB      1150 GAATAGTGATGCGCGCAGCGAGTGTGCTTCCCGCGCATACGGGATATATACCGCG 1091
OY      421 ccacatagcagaacttaaaagtgtcattatgtgaaacagttcttcggtggggaacac 480
DB      1090 CCACATAGCAGAACTTAAAGTGTCTATCATTTGAGAAAGTTCCTGCGGCGCAAAACATC 1031
OY      481 tcaagatcttaccgcgtgttgaagatccagctcagtgtaaccactcgttgcacacatga 540
DB      1030 TCAAGATCTTACCGCCTGTGTGAGATCCAGTTCAGTGAACCCATCGTGACCACTGA 971
OY      541 tctcagatcttacttacttaccacgcttctcgtgtgtgagcaaaaacaggaagcaaat 600
DB      970 TCTTCAGATCTTTTACTTTCACACGCGTTCTGGGAGACCAAAACAGGAAGCAAAAT 911
OY      601 gccgcaaaaagggaataagcgagcagcaggaatgttgaataactatactacttctctt 660
DB      910 GCCGCAAAAAGGAAATAGGCGCACGACGGAATGTTGATACATACATCTCTCTTTT 851
OY      661 caaatattgaagcattatccaggttatgtctcattgagcagatataattgaatgt 720
DB      850 CAATTTATTGAGATTTATCAGGTTATGTCTCATAGCGGATCATATTTGAATGT 791
OY      721 attagaanaataaacaataggggttccgcgcacattcccgaaaagtggcaactgac 780
DB      790 ATTTAGAAAATAAACAATAAGGGGTTCCGGCAATTTCCCGAAAAGTGCCACCTGAC 731
OY      781 gttcaagaacatattatatactatgaacttaaacctataaataagcgtlatalcagagccc 840
DB      730 GTCTAAGAAACATTTATATCATGACATTAACTTAATAAATAGCGGATCAGAGGCC 671
OY      841 ttctgcctcgcgcttctcgttgatgaacggttgaacaccccttgcacatgcagctccggag 900
DB      670 TTTCGCTCGCGGCTTTCGGTGATGACGTTGAACAACTTCATCAGATCCAGCTCCGAG 611
OY      901 acggtacagcctgtctgtatgaacgagatgcgcggaacgaagcccggtcagcgcgctca 960
DB      610 ACGGTACAGCTTGTCTGTAAAGCGGATGCGGAGACAGACAGCCCGTACAGGCGCGTCA 551
OY      961 gggggtgtggcggtgtcggggtcgtgactaactatgcggc 1001
DB      550 GCGGGTGTGGCGGCTGTGCGGGCTGCTTAATGATGCGGC 510

```

RESULT 10  
SYNPU9V/c SYNPU9V 2665 bp DNA circular SYN 26-JUL-1993  
LOCUS PUC9 Cloning vector.

ACCESSION L09128  
 VERSION L09128.1 GI:310832  
 KEYWORDS  
 SOURCE Synthetic construct  
 ORGANISM Synthetic construct  
 REFERENCE 1 (bases 1 to 2665)  
 AUTHORS Gilbert W.  
 TITLE Obtained from Vecbase 3.0  
 JOURNAL Unpublished (1991)  
 COMMENT These data and their annotation were supplied to GenBank by Will Gilbert under the auspices of the GenBank Curator Program. pUC9 - Cloning vector

ENTRY PUC9 #TYPE DNA CIRCULAR  
 DATE 17-SEP-1986  
 #sequence 16-DEC-1986  
 ACCESSION VB0022  
 SOURCE artificial  
 COLLECTION ATCC 37252  
 REFERENCE

#number 1  
 #authors Vreita J., Messing J.  
 #journal Gene (1982) 19: 259-268  
 #comment see also 'note added in proof'

REFERENCE  
 #number 2  
 #authors Pouwels P.H., Enger-Valk B.E., Brammar W.J.  
 #book Cloning Vectors, Elsevier 1985 and supplements  
 #comment vector I-A-1v-20

COMMENT  
 Assembled from pUC19 and M13mp9 by F. Pfeiffer, MPI, Martinsried  
 Revised 16-DEC-1986 by F. Pfeiffer:  
 2287/8 'Ar' to 'TA' to match revised sequence of pBR322  
 The strand shown contains the beta-galactosidase mRNA sequence  
 including the multiple cloning site of M13mp9.

## KEYWORDS

## CROSSREFERENCE

#complement

Vecbase(3):pUC9c

#parent

Vecbase(4):pUC7, Vecbase(3):M13mp9, VecSource(3):Dgal19

#brother

Vecbase(3):pUC8

#offspring

Vecbase(3):pUC13, Vecbase(3):pMBL19m, Vecbase(3):pMBL19p,

Vecbase(3):pUC931, Vecbase(3):pOM1, Vecbase(3):pOM3,

Vecbase(3):pOM9

Vecbase(3):pUC91, Vecbase(3):pUC92

Vecbase(3):pIC19H, Vecbase(3):pIC19H

Vecbase(3):pICEM19Hm, Vecbase(3):pICEM19Rp,

Vecbase(3):pICEM19Rm, Vecbase(3):pICEM19Rp

PARENT

Features of pUC9 (2665 bp)

residue source  
 1- 426 6372-6797 M13mp9  
 1- 230 1069-1298 Lac-Operon  
 233- 268 1- 36 polylinker of M13mp9  
 269- 426 1304-1461 Lac-Operon  
 427- 526 2351-2252 (C) pBR322  
 527- 663 2210-2074 (C) pBR322  
 664-2665 4355-2354 (C) pBR322  
 Conflict (cfl) and Mutations (mut):

pUC9 source  
 mut 1107 T C 3912 (C) pBR322  
 mut 1408 A G 3611 (C) pBR322

## FEATURE

931-1719 1-789 Ap-R: b-lactamase  
 POLYLINKER HindIII-PstI-SalI-BamHI-SmaI-EcoRI  
 SELECTION

## #resistance

Ap

## #indicator

beta-galactosidase

## SUMMARY

pUC9

#length 2665

#checksum 2385

FEATURES  
 source  
 Location/Qualifiers  
 1..2665  
 /organism="synthetic construct"  
 /db\_xref="taxon:32630"  
 BASE COUNT 654 a 679 c 671 g 661 t  
 ORIGIN

Query Match 100.0%; Score 1001; DB 12; Length 2665;  
 Best Local Similarity 100.0%; Pred. No. 1,2e-274;  
 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 cctgcaacttattccgcctccatccagtcattatattgttcgcggaagctagtagt 60  
 1510 CCTGCAACTTATTCGCCCTCCATCCAGTATTAATTGTCGGGAACCTGAGTAA 1451  
 61 agtcgcaagtaataatttgcaacgttggtccattgtcagagcactgtgtgtca 120  
 1450 AGTCGCAAGTAAATTAAGTTTGCCCAACGTTGTTGCCATTGTACAGGATCGTGTGCA 1391  
 121 cgtcgtcgttctgtatgcttcattcaagctccggtcccaacagatcaagagca 180  
 1390 CGTCGTCGTTGTGTGTTGCTTCACTCCGCTCCCAACGATCAAGGAGACTTACA 1331  
 181 tgatcccccagtgtgtgcaaaaagaggttaactccttcgtccctccgactgtgtca 240  
 1330 TGATCCCCCATGTTGTGCAAAAAGGAGGTACTCTCTCGGTCCTCGATGTTGTGCA 1271  
 241 agtaagtgtgcaagctgtatcaactcattgtatgagcagcactgataatccttact 300  
 1270 AGTAAGTTGGCCGCAAGTTTACATCATGCTTATGCGAGCATGCTCAATTTCTTACT 1211  
 301 gtaatccatccgtaagatgcttctgtgactgtgtgactgactcaacaaactctga 360  
 1210 GTATCCATCCGTAAGATGCTTCTGTGACTGTGACTGTAACCAAGTCAATCTGA 1151  
 361 gaatagtatgctgcaagcagctgtccttgcgagcgagtaagagataaacgag 420  
 1150 GAATAGTATGCGGCGGAGCGAGTGTGCTTGTCCCGGCTGAATACGGATATACCGCG 1091  
 421 ccacataagagaacttttaaaagtctcaatctgaaacagcttctcgggagaaaac 480  
 1090 CCACATAGCAGAACTTTAAAGTCTCATTTGGAACCACTTCTTCGGGCGCAAACTC 1031  
 481 tcaagatcttaccgctgttgatccagctcgaatgaaccactgtgacccactga 540  
 1030 TCAAGATCTTACCGCTGTGAGATCCAGTTCGATGTAACCACTCTGCAACCACTGA 971  
 541 tcttcaagcatcttactcttcaacagcgttctggtgagcaaaaagagaagcaaat 600  
 970 TCTTCAAGCATCTTACTTCAACAGCGTTCTGTGAGCAAAAACAGGAGCAAAAT 911  
 601 gccgcaaaaaggaataagagcgacacggaatggtgaatcattactctcttct 660  
 910 GCCGCAAAAAGGATATAGGCGCACACGGAATGTTGAATACATCTCTCTTTT 851  
 661 caataattgaagcaattatcaaggttatgtctcagcagagatattgaagt 720  
 850 CAATAATTGAAGCAATTATCAGGTTATGTCGTATGAGCGGATACATATTGAAAT 791  
 721 attagaataaataaacaataagaggttcgcgcgacatctcccgaaagtgcacactgac 780  
 790 ATTAGAATAAATAAACAATAAGGGTTCGCCGACATTTTCCCGAATAAGTCCACCTGAC 731  
 781 gctcaagaacacattatcatcacatcaactaataaataagaggtatccagggccc 840  
 730 GCTCAAGAACCATTTATTCATCAGCATTTAATTAATAAATAGGGATACAGAGGCC 671  
 841 ttctcgtcgcgcgttctcgttgaagcgttgaacaccttgcacatgcaagctccggag 900  
 670 TTCTCGTCCGCGGTTTCTGATGACGGTGAACCTCTGACATGAGCTCCGGAG 611  
 901 acgtcacagctgtctgtgaagcgtatgcggagagcagaacagcccgtaaggcgctca 960

Db 610 ACCGTCACAGCTTGTCTGAAGCGATGCCGGAGACAGACCCGTCAGGGCGCTCA 551  
 QY 961 ggcgggtgttggcggtgtcggagctgactactatgcgc 1001  
 Db 550 GCGGGTGTGGCGGCTGTGGGGCTGCTTAACATATGCGCG 510

RESULT 11  
 SYNPICTV/c 2668 bp DNA circular SYN 26-JUL-1993  
 LOCUS DEFINITION PIC7 cloning vector  
 ACCESSION 108880  
 VERSION 108880.1 GI:310785  
 KEYWORDS  
 SOURCE Synthetic construct DNA.  
 ORGANISM synthetic construct  
 REFERENCE 1 (bases 1 to 2668)  
 AUTHORS Gilbert, W.  
 TITLE Obtained from Vecbase 3.0  
 JOURNAL Unpublished (1991)  
 COMMENT These data and their annotation were supplied to Genbank by Will Gilbert under the auspices of the Genbank Curator Program. PIC7 - Cloning vector

ENTRY PIC7 #TYPE DNA CIRCULAR  
 TITLE PIC7 - Cloning vector  
 DATE 25-FEB-1987  
 ACCESSION V80090  
 SOURCE artificial

REFERENCE 1  
 #number 1  
 #authors Marsh J.L., Erfle M., Wykes E.J.  
 #journal Gene (1984) 32: 481-485  
 #title  
 The PIC plasmid and phage vectors with versatile cloning sites for recombinant selection by insertional inactivation

COMMENT Assembled from pUC8 and Genbank:PIC7 by F. Pfeiffer  
 COMMENT For construction of PIC7, a synthetic oligonucleotide has been used to replace the pUC8 polylinker and thus to construct a new cloning vector with a different polylinker. The other PIC-vectors are based on this new PIC7 polylinker, which was combined with the existing pUC9 and pUC19 polylinkers in the following arrangements:  
 PIC19 and PICEM19 vectors:  
 PIC19: EcoRI - Poly (PIC7) - HindIII - Poly (PUC9) - EcoRI  
 PIC19H: HindIII - Poly (PUC9) - EcoRI - Poly (PIC7) - HindIII  
 PIC20-vectors:  
 PIC20R: EcoRI - Poly (PIC7) - HindIII - Poly (PUC19) - EcoRI  
 PIC20H: HindIII - Poly (PUC19) - EcoRI - Poly (PIC7) - HindIII

COMMENT from Genbank:  
 To produce greater versatility of insertional inactivation of beta-galactosidase activity for subcloning and sequencing, a chemically synthesized oligonucleotide, specifying nine restriction sites including BglII, XhoI, NruI, ClaI, SacI and EcoI in various configurations with existing polylinkers, was created. These improved polylinkers were inserted into plasmids for routine cloning of ds-DNA and into chimeric phage/plasmids for biological production of ss-DNA. The most versatile polyrecognition pattern specifies 17 restriction sites in the beta-galactosidase alpha-complementing gene fragment. Clone PIC7 was used to produce all the other polylinker-carrying vectors.

KEYWORDS  
 CROSSREFERENCE  
 #parent Vecbase(3):PUC8, Genbank(50):PIC7  
 #offspring Vecbase(3):PIC19H, Vecbase(3):PIC19R, Vecbase(3):PIC20H, Vecbase(3):PIC20R, Vecbase(3):PICEM19Hm, Vecbase(3):PICEM19Hp,

Vecbase(3):PICEM19Rm, Vecbase(3):PICEM19Rp  
 PARENT  
 Features of PIC7 (2668 bp)  
 residue source PUC8  
 1-235 1-235  
 230-268 1-39 polylinker of PIC7  
 263-2668 260-2665 PUC8  
 Conflict (cfl) and Mutations (mut): none  
 FEATURE  
 934-1722 1-789 Ap-R; b-lactamase  
 POLYLINKER EcoRI-ClaI-EcoRV-XbaI-BglII-XhoI-SacI-NruI-HindIII  
 SELECTION  
 #resistance Ap  
 #indicator beta-galactosidase  
 SUMMARY PIC7 #length 2668 #checksum 6564.  
 Location/Qualifiers  
 1..2668  
 /organism="synthetic construct"  
 /db\_xref="taxon:32630"  
 BASE COUNT 658 a 677 c 668 g 665 t  
 ORIGIN

Query Match 100.0%; Score 1001; DB 12; Length 2668;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-274;  
 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cctgcacattatccgcctccatccatctatctatctgttcgcggaactagatgaag 60  
 Db 1513 CCTGCACATTATCCGCTCCATCCATCCAGTCACTTAATTTGTCGGGAGACCTAGATGAAC 1454  
 QY 61 agtcgcagltatagtttcgcgaacgtgtgtccattgtctacaggaactcgttgtca 120  
 Db 1453 AGTTCGCAGTAAATAGTTTGGCCAGCACTGTTGCCATTCCTACAGGCACTCGTGTGTA 1394  
 QY 121 cgttcgctgtgtatagcttcatctagctccggtcccaacagatcaagcgagtaca 180  
 Db 1393 CGCTCGTGTGTGTATGCTTCAATCACTCCGAGTCCCAAGATCAGCGAGTTACA 1334  
 QY 181 tgaaccacatgttgtcaaaaagcggttagctccctcgctccgatcgttgtcaga 240  
 Db 1333 TGATCCCATGTTGTGTGCAAAAAGCGTTAGCTCTTCGCTCCGATCGTTGTCA 1274  
 QY 241 agtaagttgcgcgaaggttcatctcatagttgtatggcaactgatatctcttact 300  
 Db 1273 AGTAAGTTGCGCGCAGTGTATCACTCATGGTATVAGCAGACAGCAFAATTTCTTACT 1214  
 QY 301 gtcagcgcacgcgaagatgtcttctcgtgactgtgtagttactcaaccagttactga 360  
 Db 1213 GTCATGCCATCCGTAAAGTCTTTTGTGTGACTGTGTAGTACTAACCAAGTCATTTCTGA 1154  
 QY 361 gaatagtatgcgcgcgcgaagttgtctctgcgcgcgcgtcaatacgggataatccgcg 420  
 Db 1153 GAATAGTGTATGCGCGCAGCAGGATGTGCTTCCCGCGGTCAATAACGGATATACCGCG 1094  
 QY 421 ccaatagcagacatttaaaagtgtcatcatctgaaacgtttctcggcgcaaaactc 480  
 Db 1093 CCACATAGCAGAACTTAAAGTGTCTATCACTTGAANAACGTTCTTTCGGCGCAAACTC 1034  
 QY 481 tcaagatcttcacgctgttgatccagttcgaatgttaaccactcgttgcaaccactga 540  
 Db 1033 TCAAGATCTTACCGCTGTGTGATCCAGTTGATGATTAACCACTCGGCACCCACTGA 974  
 QY 541 tcttaagcatcttacttcaacttcaacgaagttctcgtgtgagaaacaggaagcaaat 600  
 Db 973 TCTTACGATCTTTTACTTTCACAGCGTTCTCGGTAGCAAAAACAGGAAGCAAAAT 914  
 QY 601 gccgcaaaaaggaataaggcgacacgaagaaatlttaataactacttctcttt 660  
 Db 913 GCCGCAAAAAGGAATAAGCGCACGCAAGAAATTTGAATATCATCACTCTCTTTT 854  
 QY 661 caatatattgaagcatctatcaagggtattgtctcatgagcgagataatatattgaatg 720  
 Db caatatattgaagcatctatcaagggtattgtctcatgagcgagataatatattgaatg 720

```

Db 853 CAATATTTAGACATTATACAGGGTTATTGTCATGACGAGATACATATTTGAATGT 794
Qy 721 attagaataaataaacaataagaggttcgcgcacatttccgcaaaagtcacacccgac 780
Db 793 ATTGAAATAAATAAACAATAAGGGTTCCGGCAGCATTTCCCGCAAAAGTCCACCTGAC 734
Qy 781 gtctaaagaacattatatactacatgaacattacataaataaagagcgatcaagagccc 840
Db 733 GTCTAAGAAACCATATATATCATGACATTAACCTATAAATAAGCGTATACAGAGCCCC 674
Qy 841 ttctgcgcgcgcgttcgcgtgacgagtgaaacacttcgacacgacgtcccgag 900
Db 673 TTTCGTCTCGCGCGCTTCGCGTACGAGCGGTAACACCTGTGACACATGACGCTCCGAG 614
Qy 901 acggtcaacagctgtcgtgaagcgaatgcgcgagacagacacgcccgtcagagcgctca 960
Db 613 ACGGTACACAGCTGTCTGTAGAGCGGATCCGGGAGCAGACAGCCGCTCAGGGCGCTCA 554
Qy 961 gcgggttgtagcggttcggtggtggtggtggtggtggtggtggtggtggtggtggt 1001
Db 553 GCGGGGTGTGGCGGGGTGTGGGGGTGTGGCTTAACATATGCGGC 513

```

```

RESULT 12
SYNPUCTV/c 2674 bp DNA circular SYN 26-JUL-1993
LOCUS PUC7 Cloning vector.
ACCESSION L08958
VERSION L08958.1 GI:310823
KEYWORDS
SOURCE Synthetic construct DNA.
ORGANISM synthetic construct
REFERENCE 1 (bases 1 to 2674)
AUTHORS Gilbert, W.
TITLE Obtained from Vecbase 3.0
JOURNAL Unpublished (1991)
COMMENT These data and their annotation were supplied to GenBank by Will
Gilbert under the auspices of the GenBank Curator Program. PUC7 -
Cloning vector
ENTRY PUC7 #TYPE DNA CIRCULAR
TITLE PUC7 - Cloning vector
DATE 17-SEP-1986
#sequence 16-DEC-1986
ACCESSION V00020
SOURCE artificial
REFERENCE
#number 1
#authors Vieira J., Messing J.
#Journal Gene (1982) 19: 259-268
#comment see also 'note added in proof'
REFERENCE
#number 2
#authors Pouwels P.H., Enger-Valk B.E., Brammar W.J.
#book Cloning Vectors, Elsevier 1985 and supplements
#comment vector I-A-iv-20
COMMENT
Assembled from pUC19 and M13mp7 by F. Pfeiffer, MPI, Martinsried
Revised 16-DEC-1986 by F. Pfeiffer:
2296/7 'AT' to 'TA' to match revised sequence of PBR322
The strand shown contains the beta-galactosidase mRNA sequence
including the multiple cloning site of M13mp7.
KEYWORDS
CROSSREFERENCE
#complement
#parent Vecbase(3):pUC7c
#parent Vecbase(3):pBR322, Vecbase(3):M13mp7,
#parent GenBank(50):Ecolac, Vecsource(3):Dgal17
#offspring
Vecbase(3):pUC8, Vecbase(3):pUC9, Vecbase(3):pBR222 PARENT
Features of pUC7 (2674 bp)
residue source

```

```

1- 435 6002-6436 M13mp7
1- 229 1069-1297 lac-Operon
230- 277 1- 48 POLYlinker of M13mp7
278- 435 1304-1461 lac-Operon
436- 535 2351-2252 (c) PBR322
536- 672 2210-2074 (c) PBR322
673- 2674 4355-2354 (c) PBR322
Conflict (cfl) and Mutations (mut):
PUC7 source
mut 1116 T C 3912 (c) PBR322
mut 1417 A G 3611 (c) PBR322
FEATURE
940-1728 1-789 Ap-R: b-lactamase
POLYLINKER EcoRI-BamHI-SalI-PstI-SalI-BamHI-EcoRI SELECTION
#resistance Ap
#indicator beta-galactosidase
SUMMARY PUC7 #length 2674 #checksum 3876.
location/Qualifiers
1. 2674
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 656 a 682 c 673 g 663 t
ORIGIN

```

```

Query Match 100.0%; Score 1001; DB 12; Length 2674;
Best Local Similarity 100.0%; Pred. No. 1.2e-274;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 cctgaacattatccgcctccatccatctataattgttcgcggaagctagaagt 60
Db 1519 CCTGAACATTATCCGCCCMCCATCCAGTCTATATTTGTCGCCGGAGCTAGAGT 1460
Qy 61 agtgcagatgaatgaattgcgcagcgtgttccattgtctaacgagcatgtgtgtca 120
Db 1455 AGTGCAGATTAAAGTTTGCGACAGCTTGGCATTTGCTACAGCATCTGTGTGCA 1400
Qy 121 cgtctgcgttggatagcttcatcttcagctccggttcccaagcatcaagcgagttaca 180
Db 1399 CGCTCGTGTGGTATGAGTTCATTCATTCAGCTCCGTTCCCAAGCATCAAGCGAGTTACA 1340
Qy 181 tgaatcccatggtgtgtgcaaaaagcggttagtgccttccttgatccgacgtgtgtcaga 240
Db 1339 TGATCCCATGTTGTGCAAAAAGCGGTAGCTTCCTGCTCCGATCTGTCAGA 1280
Qy 241 agtaagttgcgcagtggtatcaactaagttatggtatgagcagctcatcttctact 300
Db 1279 AGTAAGTTGGCGCGAGTGTATCATCTCATGTTATGCGACGACTGCATTAATTCCTTACT 1220
Qy 301 gtcattgcattcgtgaagatgtcttctgtgactgtgtgagttactcaaccagatcttga 360
Db 1219 GTCATGCCATCCGTAAGATGCTTTGTGTGACTGTGATGATCAACCAAGTCAATCTGA 1160
Qy 361 gaatagttatgagcgagcagagttgtcttgcgcgcgcgtcgaatagaggtatccgg 420
Db 1159 GAATAGTTATGCGCGACCGAGTGTGCTTGTGCGCGCTGATACCGCGGATTAACCGCG 1100
Qy 421 ccacatagcagaacttaaaagtgctcatcatctgaaacagctcttcgggagcaaac 480
Db 1099 CCACATAGCAGAACTTAAAGGTCATCATTTGGAAAGAGTTCTTCGGGCGCAAACTC 1040
Qy 481 tcaaggtcttaccgctgtgtgagatccagttcagttgaaccacatcgtgacccaactga 540
Db 1039 TCAAGGATCTTACCGCGTGTGAGATCCAGTTCATGTAACCCACTGTCGACCACTGA 980
Qy 541 tcttcaacatcttacttaccacagaggttctgggtgtgagcaaaaacagaggaacaaat 600
Db 979 TCTTCACATCTTTTACTTTCACACAGGTTTCTGGGTGACCAAAACAGGAAGCAAAAT 920
Qy 601 gcgcgcaaaaaggaataagagcgacacggaataatgttgatactactactcttctt 660
Db 919 GCGCAAAAAGGAATTAAGGCGACACGGAATGTTGAATCTACTCTTCTTTT 860

```





## COMMENT

These data and their annotation were supplied to Genbank by Will Gilbert under the auspices of the Genbank Curator Program. pUC13 - Cloning vector

## #TYPE DNA CIRCULAR

TITLE pUC13 - Cloning vector

DATE 17-SEP-1986

#sequence 16-DEC-1986

ACCESSION V00024

SOURCE artificial

## REFERENCE

#number 1

#authors Vieira J., Messing J.

#journal Gene (1982) 19: 259-268

#comment see 'note added in proof'

## REFERENCE

#number 2

#authors Pouwels P.H., Enger-Valk B.E., Brammar W.J.

#book Cloning Vectors, Elsevier 1985 and supplements

#comment vector I-A-iv-20

## COMMENT

Assembled from pUC19 and M13mp11 by F. Pfeiffer, MPI, Martinsried

Revised 16-DEC-1986 by F. Pfeiffer:

2302/3 'AT' to 'TA' to match revised sequence of PBR322

The strand shown contains the beta-galactosidase mRNA sequence

including the multiple cloning site of M13mp11. KEYWORDS

## CROSSREFERENCE

#complement

Vecbase(3):pUC13c

#parent Vecbase(3):pUC9, Vecbase(3):M13mp11, VecSource(3):bcall3

#brother Vecbase(3):pUC12

#offspring Vecbase(3):pUC19, Vecbase(3):pUC19c,

Vecbase(3):pSP65, Vecbase(3):pGEM2, Vecbase(3):pT713 PARENT

Features of pUC13 (2680 bp)

1-441 6002-6442 M13mp11

1-230 1069-1298 Lac-Operon

233-283 1-51 polylinker of M13mp11

284-441 1304-1461 Lac-Operon

442-541 2351-2252 (c) PBR322

542-678 2210-2074 (c) PBR322

679-2680 4355-2354 (c) PBR322

Conflict (cfl) and Mutations (mut):

Vecbase(3):pUC13c

mut 1122 T C 3912 (c) PBR322

mut 1423 A G 3611 (c) PBR322

FEATURE

946-1734 1-789 Ap-R: b-lactamase

POLYLINKER HindIII-PstI-SalI-XbaI-BamHI-SmaI-SacI-EcoRI SELECTION

#resistance Ap

#indicator beta-galactosidase

SUMMARY pUC13 #length 2680 #checksum 217.

location/Qualifiers

1..2680

/organism="Cloning vector pUC13"

/db\_xref="taxon:154051"

misc\_feature

1..441

/note="M13mp11"

misc\_feature

1..230

/note="Lac-Operon"

misc\_feature

233..283

/note="M13mp11 polylinker"

misc\_feature

284..441

/note="Lac-Operon"

misc\_feature

442..2680

/note="PBR322"

CDS

879..1739

/note="Ap-R"

/codon\_start=1

/transl\_table=11

/product="beta-lactamase"  
/protein\_id="AAK16201.1"  
/db\_xref="GI:13236870"  
/translation="MSTQHRVALIPFAFCIPVAPHETLYKVKDAEPLGARYC  
IEDLNSGKILIESFRPEERPMSTFVLLCGAVLSIDGQDLGRHITSONDLVE  
YSPVTEKHLDGMYRELCSAATMSNTANLTTTGGPKELTFLHMGHVRLL  
DRMEPELNEAIPINDERDTPVAMATLRLGELLTTLASROOLDMMKADKVASPL  
LRSAIPAGWFIADSGAGERSGRIIALAPDPKPSRIVLYTTGSQATMDERNROIA  
EIGSLIKHW"  
1122  
/replac="c"  
variation  
1423  
/replac="g"  
BASE COUNT 657 a 683 c 676 g 664 t  
ORIGIN

Query Match 100.0%; Score 1001; DB 12; Length 2680;  
Best Local Similarity 100.0%; Pred. No. 1.2e-274;  
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 cctgaacttattccgcctccatcagctctattgttgcgggaagctagaagt 60  
|||||  
Db 1525 CCTGCACTTATMCCGCCATCCAGCTATTAATGTGCGGGAAGTAGTAAGT 1466  
61 agtcgcaatgaattgttgcgaagcttgcattgttgcagcagtcgtgtgca 120  
|||||  
Db 1465 AGTTGCCAGTTAATAGTTGGCGCAACGTTGGCATTTGTCAGAGCATGTTGTC 1406  
121 cgtcgtcgttctgtatagctcattcagctcgcgttcccaagatcaagcag 180  
|||||  
Db 1405 CGCTCGTCTTGGTATGAGCTTCATTCAGCTCCGCTCCCAAGATCAAGGAGTACA 1346  
181 tgatcccccattgttgcgaagaaacggttaagctcttccttcggtcccgatg 240  
|||||  
Db 1345 TCATCCCATGTTGTGCAAAAACGGTTACTCTTCGGCTCCGATCGTTCAGA 1286  
241 agtaagttgacgacgtgttctacatcagctgttgcagcagcagtcataatct 300  
|||||  
Db 1285 AGTAAGTTGGCGGAGGCTTATCAGCATGTTATGGCAGCATCAATAATCTTACT 1226  
301 gtcagtcacatcgttaagatgcttctgtgactgtgtgactcaacaaatcatt 360  
|||||  
Db 1225 GTCATGCCATTCGTAAGATGCTTCTTGACTGGTGAATCTCAACCAAGTCATTTGA 1166  
361 gaatagtatgacgacgacgagttgctcttgcggcgctcaatacggatatacc 420  
|||||  
Db 1165 GAATAGTGTATGCGGCGAGGAGTTGCTTGGCCGCGCTCAATACGGATATACCGG 1106  
421 ccacatagcagaactttaaaagtctcattcattgaaacgttcttcgggacgaact 480  
|||||  
Db 1105 CCACATAGCAGAACTTTAAAGAGCTCATCATTTGAAACGTTCTTGGCGGGAACATC 1046  
481 tcaagatcttaccgctgtgtgagatcagttgagtaacccacgctgacccaaact 540  
|||||  
Db 1045 TCAAGAGATCTTACCGCTGTGAGATCAGTTGAGTAAACCACTGCTGACCACTGA 986  
541 tcttcagcatcttcttaccctcaccagcgttcttcgggtgagcaaaaacaggaag 600  
|||||  
Db 985 TCTTCAGCATCTTTTACTTTCAACACGCTTTCGGGTGAGCAAAAACAGGAAGCA 926  
601 gcccaaaaaaggaaatgaaggcgacacggaatgttgatctcattcttctt 660  
|||||  
Db 925 GCCCAAAAAAGGAAATGAGGCGACACGGAATGTGAAATCTCTACTCTTCTTTT 866  
661 caatatattgaagcattatagggttatgtctcaatgaagcggaataattgaagt 720  
|||||  
Db 865 CAATATTATTGAAGCATTTATCAGGGTATGTCTATAGCGGATACATATTGAATGT 806  
721 attagaaaataaacaataggggttcgcgcacattcccgaaaagtgcacactgac 780  
|||||  
Db 805 ATTGAAAAAATTAACAAATAGGGGTTCCGGCGACATTTCCCGAAAGTGCACCTGAC 746

Oy 781 gtctaagaacacattatcatcatgaacctaataaalaagcgylatcacgagggcc 840  
|||||  
Db 745 gtcTAgAAACCATATATATCATGACATTAACTATAAAATAGGCGTATCAGAGGCC 686  
|||||  
Oy 841 ttgcgtcgcgcggtttcggtgatgaagtgaaacctctgacacatgcaagctcccgag 900  
|||||  
Db 685 TTTCGTCTGCGCGGTTTGCGGTGATGACGGTGAAAACTGTGACACATGACAGCTCCGGAG 626  
|||||  
Oy 901 acggtcaacagcttctgtaagcgatgacgagcagacaagccgltcaagggcgctca 960  
|||||  
Db 625 ACGGTCAACAGCTTGTCTCTGAGCGGATCCCGGAGCAGACAGCCCGTCAAGGCGCGTCA 566  
|||||  
Oy 961 gcggtgttgcgaggggtgcgggcttgcttaactatgcgagc 1001  
|||||  
Db 565 GCGGCTTTGGCGGGGTGCGGGCTGAGCTTAACATATGCGGC 525  
|||||

Search completed: January 8, 2002, 17:16:51.  
Job time: 12102 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 8, 2002, 17:26:36 ; Search time 428.83 Seconds  
(without alignments)  
2001.219 Million cell updates/sec

Title: US-09-740-211-13\_COPY\_6000\_7000

Perfect score: 1001

Sequence: 1 cctgcaacttaccgcctc.....ggctgcttaactatgcgcgc 1001

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_1101.\*

- 1: /SIDS2/gcgdata/geneseq/geneseqn/NA1980.DAT.\*
- 2: /SIDS2/gcgdata/geneseq/geneseqn/NA1981.DAT.\*
- 3: /SIDS2/gcgdata/geneseq/geneseqn/NA1982.DAT.\*
- 4: /SIDS2/gcgdata/geneseq/geneseqn/NA1983.DAT.\*
- 5: /SIDS2/gcgdata/geneseq/geneseqn/NA1984.DAT.\*
- 6: /SIDS2/gcgdata/geneseq/geneseqn/NA1985.DAT.\*
- 7: /SIDS2/gcgdata/geneseq/geneseqn/NA1986.DAT.\*
- 8: /SIDS2/gcgdata/geneseq/geneseqn/NA1987.DAT.\*
- 9: /SIDS2/gcgdata/geneseq/geneseqn/NA1988.DAT.\*
- 10: /SIDS2/gcgdata/geneseq/geneseqn/NA1989.DAT.\*
- 11: /SIDS2/gcgdata/geneseq/geneseqn/NA1990.DAT.\*
- 12: /SIDS2/gcgdata/geneseq/geneseqn/NA1991.DAT.\*
- 13: /SIDS2/gcgdata/geneseq/geneseqn/NA1992.DAT.\*
- 14: /SIDS2/gcgdata/geneseq/geneseqn/NA1993.DAT.\*
- 15: /SIDS2/gcgdata/geneseq/geneseqn/NA1994.DAT.\*
- 16: /SIDS2/gcgdata/geneseq/geneseqn/NA1995.DAT.\*
- 17: /SIDS2/gcgdata/geneseq/geneseqn/NA1996.DAT.\*
- 18: /SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT.\*
- 19: /SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT.\*
- 20: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT.\*
- 21: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT.\*
- 22: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1001	100.0	2462	21	AAAT74638	Plasmid pSP72. Es
2	1001	100.0	2686	22	AAAT59053	Plasmid vector pUC
3	1001	100.0	3190	19	AAAT32980	Tn7 target plasmid
4	1001	100.0	3253	19	AAAT34339	DNA sequence of th
5	1001	100.0	3331	21	AAAT92245	Nucleotide sequenc
6	1001	100.0	3343	11	AAAT04655	Plasmid p10169 enc
7	1001	100.0	3383	21	AAAT51632	Plasmid pGM678 con
8	1001	100.0	3444	21	AAAT51630	Plasmid pGM679 con
9	1001	100.0	3619	18	AAAT78825	Kappa light chain
10	1001	100.0	3619	19	AAAT39266	Plasmid pUC65 nuc
11	1001	100.0	3819	20	AAAT22020	Nucleotide sequenc

C	12	1001	100.0	3858	22	AAAT04947	Plasmid pRK50 used
	13	1001	100.0	3880	19	AAAT39242	Plasmid pCK7-96 nu
	14	1001	100.0	3881	18	AAAT78801	Kappa light chain
	15	1001	100.0	3881	20	AAAT21996	Nucleotide sequenc
	16	1001	100.0	3903	22	AAAT6078	Feline IL-16 encod
	17	1001	100.0	3938	20	AAAT61041	Plasmid pPRH43 con
	18	1001	100.0	4009	19	AAAT00679	Plasmid pND211 con
	19	1001	100.0	4045	15	AAAT00942	Plasmid pTR3NMI co
	20	1001	100.0	4074	21	AAAT48267	Construct pGEM-hTR
	21	1001	100.0	4118	18	AAAT6186	Construct pGEM-hTR
	22	1001	100.0	4118	18	AAAT6189	Plasmid pGM712 con
	23	1001	100.0	4189	21	AAAT51634	Feline IL-12p35 en
	24	1001	100.0	4201	22	AAAT6077	DNA sequence of he
	25	1001	100.0	4421	17	AAAT58319	Feline IL-12p40 en
	26	1001	100.0	4522	22	AAAT6076	Plasmid pMO1-f1112
	27	1001	100.0	4522	22	AAAT6076	altR reading frame
	28	1001	100.0	4554	21	AAAT58319	Plasmid pSUN387.
	29	1001	100.0	4603	11	AAAT04010	DNA sequence of p1
	30	1001	100.0	4613	22	AAAT59062	Plasmid pUCM5-6h73
	31	1001	100.0	4618	21	AAAT39628	Nucleotide sequenc
	32	1001	100.0	4694	22	AAAT55225	Murine Iad alpha c
	33	1001	100.0	4713	19	AAAT12067	Gamma heavy chain
	34	1001	100.0	4723	18	AAAT78802	Plasmid pCG7-96 nu
	35	1001	100.0	4723	19	AAAT39243	Nucleotide sequenc
	36	1001	100.0	4723	20	AAAT21997	Nucleotide sequenc
	37	1001	100.0	4723	22	AAAT55224	Murine Iad beta ch
	38	1001	100.0	4724	19	AAAT12068	Expression constru
	39	1001	100.0	4776	20	AAAT7617	Plasmid pHC65 nuc
	40	1001	100.0	4776	20	AAAT7614	Nucleotide sequenc
	41	1001	100.0	4926	19	AAAT39291	Retroviral vector
	42	1001	100.0	4926	20	AAAT22045	Transdominant effe
	43	1001	100.0	4950	18	AAAT03801	pcMV-1-Cre-PA vect
	44	1001	100.0	4950	22	AAAT82936	
	45	1001	100.0	4960	22	AAAT09269	

## ALIGNMENTS

RESULT 1

ID AAAAT74638 standard; DNA: 2462 bp.

XX AAAAT74638;

XX 05-DEC-2000 (first entry)

XX DT

XX DE Plasmid pSP72.

XX XX Plasmid pSP72; NSC-p25 transgene; p25; cdk5 activator;

KW neuron specific enolase; NSC; promoter; neurodegenerative disease;

KW Alzheimer's disease; Parkinson's disease; amyloid lateral sclerosis;

KW Huntington's disease; traumatic brain injury; stroke; transgenic animal;

KW spinocerebellar degeneration; tau hyperphosphorylation; ds.

XX OS Escherichia coli.

XX PN EP1026251-A2.

XX PD 09-AUG-2000.

XX PF 02-FEB-2000; 2000EP-0300797.

XX PR 03-FEB-1999; 99US-0118478.

XX PA (PFIZ ) PFIZER PROD INC.

XX PT AbLifanin MK, Mcneish JD;

XX DR WPI: 2000-507252/46.

XX PT New recombinant DNA molecule comprising gene encoding p25, useful for producing transgenic animal to serve as model for



Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 cctgcaacttaccgcctccacacagctcattatgttccggaagctagaagta 60
    |||
DB 1531 CCTGCAACTTATCCGGCTCCATCCAGTATTAATGTGCGGGAGCTAGTAAGT 1472

OY 61 agtcgcagtaataagttgacgaacggttgctgcaattgctacagcgtggtgta 120
    |||
DB 1471 AGTTCGCCAGTTAATAGTTTGGCAACGTTGTGCGCATTCCTACGACATCGGTGTGA 1412

OY 121 cgtcgcgtcttggtgatggttcatcagctccggttcccaacagatcaagcagttaca 180
    |||
DB 1411 CGCTCGTGTGTTGTTAGGCTTCAATTCAGCTCCGGTTCACAGATCAAGCGAGTTACA 1352

OY 181 tgaatcccaatgttctgcaaaaaaagcggttagctccttcggtccctccgactgtgtcaga 240
    |||
DB 1351 TGATCCCCCATGTGTSCAAAAAAGCGTTAGCTCCTCGTCCGATCGTGTGTCAGA 1292

OY 241 agtaagttgcccgcagtggttatacctcatgtgtatagcagcactgcaatacttact 300
    |||
DB 1291 AGTAAGTTGGCCGACGTTATACACTCATGTGTATGGACACACTGCAATTAATCTCTTACT 1232

OY 301 gtcatgcatcgcgaagatgcttctctgtactgtgtagtaactcaacaaagtaacttga 360
    |||
DB 1231 GTCATGCCATCCGTAAGATGCTTTTCTGTGACTGTGATGTAACCAAGTCAATTTCTGA 1172

OY 361 gaatagttatgacgacgacgagttgtcttccggtcgctcaatacagagataaccgag 420
    |||
DB 1171 GAATAGTGTATGCGGCGACGAGTGTGCTTCCGCGCGTCAATACGCGAATAACCGCG 1112

OY 421 ccaatagcagaactttaaaagtgctcatatgtgaaacggttcttcggaacgaacatc 480
    |||
DB 1111 CCACATACCAACACTTAAAGTCTCATGTGAAACGTTCTTGCGGGCGAAACATC 1052

OY 481 tcaagatcttaccgctgtctgataccagttcga tgaacccactcgtgcacccaactga 540
    |||
DB 1051 TCAAGATCTTACCGCTGTGTAGATCCAGTTCGATGTAAACCATCTGTCACCCAACTGA 992

OY 541 tcttcagcatcttacttctcaacagcgttcttggtgagagaaaaaaggaagcaaat 600
    |||
DB 991 TCTTCAGCATCTTWTACTTCTCAAGCGTCTTGCGGTGAGCAAAACAGGAAGCAAAAT 932

OY 601 gccgcaaaaaaggaataaagggcgaacgaaatgttgaatactactacttctctttt 660
    |||
DB 931 GCCGCAAAAAAGGGAATAGGGCGACAGGAATGTTCAATACATCACTCTCTCTTTT 872

OY 661 caataattgaagcattatcaaggttatctgtctcaatgacgagataatattgaaatgt 720
    |||
DB 871 CAATATTTATTGAAAGCATTTATCAGGTTATGTTCTCATGACGCGATACATATTGAAATGT 812

OY 721 atttagaaaataaacaataaggggttccgcgcacatttcccgaaagtgccacctgac 780
    |||
DB 811 ATTTAGAAAATTAACAAATAGGGGTTCCGCGCACATTTCCCGAAAGTGCACACTGAC 752

OY 781 gttcaagaacacatla b c t c a t g a c a l t a a a a t a a g g c g l a c a g a g c c c 840
    |||
DB 751 GCTTAAGAAACCATATTATGATGATTAACCTTAATAAATAGCGGTATACAGAGGCC 692

OY 841 ttctgtctcgcggttctggtgtatgacggtgaaacccctgaacacatgcaatcccgag 900
    |||
DB 691 TTTTCGCTTCGCGGTTTGGTGTATGACGGTGAACCTCTGTGACATGACAGTCCCGAG 632

OY 901 acggtacacgctgtctgtgaagcgaatgcggaagcagacagaacagcccgctcagggcggtca 960
    |||
DB 631 ACAGGTACAGCTGTGCTGTATACCGGATGCCGGAGCAGACAGACCCGCTACAGGCGCGTCA 572

OY 961 gcgggtgttgccggtgtgtcggggtgctgacttaactatgagc 1001
    |||
DB 571 GCGGGTGTGCGGGGTGCGGGCTGAGCTTAACATATGCGCG 531

```

RESULT 3  
AAV32980/c

ID AAV32980 standard; DNA: 3190 BP.

```

XX AC AAV32980;
XX XX
XX 17-NOV-1998 (first entry)
XX XX
XX Tn7 target plasmid sequence.
XX DE
XX Tn7 target plasmid; transposon; ATP-utilising regulatory protein;
XX KW transposable element; DNA sequencing; genetic analysis;
XX KW insertional mutagenesis; ss.
XX OS
XX Synthetic.
XX PN W09837205-A1.
XX PD 27-AUG-1998.
XX XX
XX PF 20-FEB-1998; 98WO-US03353.
XX PR 20-FEB-1997; 97US-0037955.
XX XX
XX (CRAI/) CRAIG N L.
XX PA
XX Craiy NL;
XX PI
XX DR WPI: 1998-467567/40.
XX XX
XX PT Transposon(s) encoding mutant ATP using proteins for insertion -
XX PT which is efficient and random, with reduced site specificity; for
XX PT DNA sequencing and altering gene expression
XX XX
XX PS Disclosure: Fig 11B; 143pp; English.
XX XX
XX CC The present sequence represents a Tn7 target plasmid pBR 183.
XX CC Transposon Tn7 encodes an ATP-utilising regulatory protein that
XX CC contains a mutation that allows efficient and simple insertion of,
XX CC and reduced target site specificity of, a transposable element derived
XX CC from the transposon. The ATP-regulatory protein and compositions are used
XX CC for the efficient, non-specific and simple insertion of a transposon or
XX CC transposable element into a DNA segment. This is useful in DNA
XX CC sequencing, for genetic analysis by insertional mutagenesis, or for
XX CC alteration of gene expression by insertion of a desired sequence.
XX XX
XX SQ Sequence 3190 BP; 784 A; 808 C; 773 G; 825 T; 0 other:

Query Match 100.0%; Score 1001; DB 19; Length 3190;
Best Local Similarity 100.0%; Pred. No. 3,7e-263;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 cctgcaacttaccgcctccacacagctcattatgttccggaagctagaagta 60
    |||
DB 2035 CCTGCAACTTATCCGGCTCCATCCAGTATTAATGTGTCGCGGAAGCTAAGTAAGT 1976

OY 61 agtcgcagtaataagttgacgaacggttgctgcaattgctacagcgtggtgta 120
    |||
DB 1975 AGTTCGCCAGTTAATAGTTTGGCAACGTTGTGCGCATTCCTACGACATCGGTGTGA 1916

OY 121 cgtcgcgtcttggtgatggttcatcagctccggttcccaacagatcaagcagttaca 180
    |||
DB 1915 CGCTCGTGTGTTGTTAGGCTTCAATTCAGCTCCGGTTCACAGATCAAGCGAGTTACA 1856

OY 181 tgaatcccaatgttctgcaaaaaaagcggttagctccttcggtccctccgactgtgtcaga 240
    |||
DB 1855 TGATCCCCCATGTGTSCAAAAAAGCGTTAGCTCCTCGTCCGATCGTGTGTCAGA 1796

OY 241 agtaagttgcccgcagtggttatacctcatgtgtatagcagcactgcaatacttact 300
    |||
DB 1795 AGTAAGTTGGCCGACGTTATACACTCATGTGTATGGACAGACTGCAATTAATCTCTTACT 1736

OY 301 gtcatgcatcgcgaagatgcttctctgtactgtgtagtaactcaacaaagtaacttga 360
    |||

```



QY 601 gccgcacaaaggaatgaggcgacacggaatgtgaatacactactctctt 660  
 Db 2133 gccgcacaaaggaatgaggcgacacggaatgtgaatacactactctctt 2192  
 QY 661 caatatattgaagcattatcaaggattatgtctcatgacgagatacatttgaatgt 720  
 Db 2193 caatatattgaagcattatcaaggattatgtctcatgacgagatacatttgaatgt 2252  
 QY 721 attagaataataaataaaggttcgcgcgcacattccccgaaagagtcacactgac 780  
 Db 2253 attagaataataaataaaggttcgcgcgcacattccccgaaagagtcacactgac 2332  
 QY 781 gcttaagaacacattatctcatgacatacctataaataagcgatcaagagccc 840  
 Db 2313 gcttaagaacacattatctcatgacatacctataaataagcgatcaagagccc 2372  
 QY 841 ttctgctcgcgcgttcggtgagtgagtgaaacactctgacacatgcagctccggag 900  
 Db 2373 ttctgctcgcgcgttcggtgagtgagtgaaacactctgacacatgcagctccggag 2432  
 QY 901 acggtcacacattgtctgtaagcaggtatgcgcgagcagaacagccgtcaagcgctgca 960  
 Db 2433 acggtcacacattgtctgtaagcaggtatgcgcgagcagaacagccgtcaagcgctgca 2492  
 QY 961 gcgggtgtgacgggtgtcgcgggctggtgcthaactatgcgac 1001  
 Db 2493 gcgggtgtgacgggtgtcgcgggctggtgcthaactatgcgac 2533

RESULT 5  
 AA299245/C  
 ID AA299245 standard; DNA: 3331 BP.

AA299245:

03-JUL-2000 (first entry)

Nucleotide sequence of a Pinpoint expression vector.

Pinpoint vector: fusion protein antigen: membrane antigen: syphilis: ss.

Treponema pallidum.

EP985931-A2.

15-MAR-2000.

12-AUG-1999; 99EP-0115877.

04-SEP-1998; 98US-0148920.

(BECT) BECTON DICKINSON & CO.

Mullenix MC, Deutsch J;

WPI: 2000-226057/20.

Immunosay for anti-Treponema pallidum antibodies, used for diagnosis of syphilis using a fusion protein of membrane antigen with peptide sequence that can be biotinylated in vivo

Disclosure: Page 8-9; 16pp; English.

The present sequence represents the Pinpoint vector, which is used in the course of the invention. The specification describes a method for detecting antibodies against Treponema pallidum. The antibodies are detected in a sample by reaction with a fusion protein antigen, present in the mixture in limiting concentration. The fusion protein antigen comprises a Treponema pallidum membrane antigen. The method is used for diagnosis of syphilis.

Sequence 3331 BP; 811 A; 854 C; 885 G; 781 T; 0 other:

Query Match 100.0%; Score 1001; DB 21; Length 3331;  
 Best Local Similarity 100.0%; Pred. No. 3,7e-263;  
 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cctgcacattatccgcgccacatccagctatattgttgcgggaagctagaatgt 60  
 Db 1565 cctgcacattatccgcgccacatccagctatattgttgcgggaagctagaatgt 1506  
 QY 61 agtcgcagtaataagattgtgcgaacgttgttgcacattgtctaaaggaatgtgtca 120  
 Db 1505 agtcgcagtaataagattgtgcgaacgttgttgcacattgtctaaaggaatgtgtca 1446  
 QY 121 cgttcgttgttgtatgtgtctcaatcaagctccggttcccaagataagcgagttca 180  
 Db 1445 cgttcgttgttgtatgtgtctcaatcaagctccggttcccaagataagcgagttca 1386  
 QY 181 tgaatcccaatgtgtgcaaaaaagcggttagctccctgcgtccctcgatgtgtcaga 240  
 Db 1385 tgaatcccaatgtgtgcaaaaaagcggttagctccctgcgtccctcgatgtgtcaga 1326  
 QY 241 agtaagttgacgcagtgatgtatcatcactatgtgtatgtgacagcactgatattcttact 300  
 Db 1325 agtaagttgacgcagtgatgtatcatcactatgtgtatgtgacagcactgatattcttact 1266  
 QY 301 gtaatgccatccgtaagatgtcttctgtgactgtgtgagtaactcaacgaatcattcga 360  
 Db 1265 gtaatgccatccgtaagatgtcttctgtgactgtgtgagtaactcaacgaatcattcga 1206  
 QY 361 gaatagttatgtgcgcagcaggttctcttgcgcggcgtcaaaaggaataaacgcgcg 420  
 Db 1205 gaatagttatgtgcgcagcaggttctcttgcgcggcgtcaaaaggaataaacgcgcg 1146  
 QY 421 ccacatagcaagaacttaaaagtgctcatcatatgtgaaacgcttctcggcgcaaacctc 480  
 Db 1145 ccacatagcaagaacttaaaagtgctcatcatatgtgaaacgcttctcggcgcaaacctc 1086  
 QY 481 tcaagatcttaacgcgtgtgtgagatcaggttgatgtgaacacactgtgtgacccactga 540  
 Db 1085 tcaagatcttaacgcgtgtgtgagatcaggttgatgtgaacacactgtgtgacccactga 1026  
 QY 541 tcttcagatcttacttacttccacagcgttctgtgtgagcaaaaacgaaggaacaaat 600  
 Db 1025 tcttcagatcttacttacttccacagcgttctgtgtgagcaaaaacgaaggaacaaat 966  
 QY 601 gccgcacaaaggaataagggcgacacggaatgttgaatactcaactcttctctt 660  
 Db 965 gccgcacaaaggaataagggcgacacggaatgttgaatactcaactcttctctt 906  
 QY 661 caatatattgaagcattatcaaggttatgttctcatgacgagatcatatttgaatgt 720  
 Db 905 caatatattgaagcattatcaaggttatgttctcatgacgagatcatatttgaatgt 846  
 QY 721 attagaataataaataaaggttcgcgcgcacattccccgaaagagtcacactgac 780  
 Db 845 attagaataataaataaaggttcgcgcgcacattccccgaaagagtcacactgac 786  
 QY 785 gcttaagaacacattatctcatgacatacctataaataagcgatcaagagccc 840  
 Db 725 gcttaagaacacattatctcatgacatacctataaataagcgatcaagagccc 840  
 QY 841 ttctgctcgcgcgttcggtgagtgagtgaaacactctgacacatgcagctccggag 900  
 Db 725 ttctgctcgcgcgttcggtgagtgagtgaaacactctgacacatgcagctccggag 900  
 QY 901 acggtcacacattgtctgtaagcaggtatgcgcgagcagaacagccgtcaagcgctgca 960  
 Db 665 acggtcacacattgtctgtaagcaggtatgcgcgagcagaacagccgtcaagcgctgca 960  
 QY 961 gcgggtgtgacgggtgtcgcgggctggtgcthaactatgcgac 1001  
 Db 605 gcgggtgtgacgggtgtcgcgggctggtgcthaactatgcgac 565

```

RESULT 6
AA004655/c
ID AA004655 standard; DNA: 3343 BP.
XX AC
XX AA004655;
XX DT
XX 01-OCT-1990 (first entry)
XX DE
XX Plasmid p10169 encoding antibody MAK33 Fd chain.
XX KW
XX Antibody MAK33; Fd chain; plasmid p10169; activated antibody prodn.; ss.
XX FH
XX Key Location/Qualifiers
XX CDS 240..917
XX FT /*tag= a
XX FT /product= Fd chain of MAK33
XX PN
XX DE3835350-A.
XX PD
XX 19-MAR-1990.
XX PF
XX 17-OCT-1988; 88DE-3835350.
XX PR
XX 17-OCT-1988; 88DE-3835350.
XX PA
XX (BOEF ) BOEHRINGER MANNHEIM GMBH.
XX PI
XX Rudolph R, Buchner J, Lenz H;
XX DR
XX WPI; 1990-156813/21.
XX PT
XX Activated antibody prodn. from recombinant procarotoc cells -
XX PT by solubilisation under reducing conditions, then oxidative
XX renaturation, carried out at low protein concn.
XX PS
XX Disclousure; ; P; German.
XX CC
XX Plasmid pBFI11 encoding the kappa chain and plasmid AAP10169 encoding
XX CC the Fd chain of antibody MAK33 were used to transform E.coli DSM 3689
XX CC and the resulting cells grown to form inclusion bodies. After the final
XX CC renaturation step an 18% yield of biologically-active protein was
XX CC produced.
XX CC see also AA004654.
XX SQ
Sequence 3343 BP: 815 A; 871 C; 846 G; 811 T; 0 other;

Query Match 100.0%; Score 1001; DB 11; Length 3343;
Best Local Similarity 100.0%; Pred. No. 3.8e-263;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 cctgaacttatacgcctcatccatcagtcataatgttgcgggaagctagaagtaagt 60
DB 2188 CCTGCAACTTTATCCGCTTCATCCAGCTATTAATTGTTGCCGGAAGCTAGATAAGT 2129
OY 61 agttcgcaatgaatagttgcgaacgttctgcaatgtctacacaggaacatcggtgtca 120
DB 2128 AGTTGCCAGTTAATAGTTTGCAGACGTTCTTCCATTCTACAGGCACTGtGtGtCA 2069
OY 121 cgtctgtcgttctgtataggtctcattcagtcgcgttcccaagcatcaagcgagttaca 180
DB 2068 CGCTGCTGTTGGTATGGCTTCAATCCGCTCCGATCCCAACGATCAAGGCGAGTTACA 2009
OY 181 tgatcccccagttgtgcaaaaagcggttagctccttgcgttcctccgatcgttgcaga 240
DB 2008 TGATCCCCCAATGTTGTGCAAAAAGCGGTTAGCTCCTTCGCTCCGATCGTTGTCACA 1949
OY 241 agtaagtgtgcgcagtgatatacactatgltatagcagcaactgataatcttact 300
DB 1948 AGTAAGTGGCCGCGAGTATATCATCATGTTATGCGAAGCACTGATTAATCTTCTTACT 1889

```

```

OY 301 gtcatgtccatccgtaaatgcttctctgtactgtgtgactactcaaccaagtcattcga 360
DB 1888 GTCATGCCATCCGTAAAGTGTCTTGTGTAGTGTAGTACTCAACCAAGTCATTCTTA 1829
OY 361 gaatagtgtatgcgcgcacgcagtgctctctccgcgcgtcgaataacacgcgcg 420
DB 1828 GAATAGTGTATGCGCGACCGAGTGTCTTGGCCGCGCATATACGGGATTAATACCGG 1769
OY 421 ccacatagcagaacttaaaagtgtcctcatcattgtgaaacgttcttcggggagaaacac 480
DB 1768 CCACATAGCAGAACTTTAAAGTGCTCATCATGTGAAAAAGCTTTCTTGGGCGGAAACATC 1709
OY 481 tcaagatcttaccgcgtgttgagatccagltcgaatgaaccactcgtgcaccaactga 540
DB 1708 TCAAGATCTTACCGCTGTGTAGATCCAGTTCGATTAACCCATCTGTCACCCAACTGA 1649
OY 541 tctcagcatcttctcattcaccagcgttcttcgggtgagaaacgcggaaggaacaaat 600
DB 1648 TCTTCACGATCTTTTACTTTCACACAGCGTTTCTTGGGTGAGCAAAAACAGAAAGCAAAAT 1589
OY 601 gcgcgaataaagggaatgaaggcgacacggaatgttgatactactcttcttctt 660
DB 1588 GCGCAAAAAGGGAATTAAGGCGGACACGAAATGTTGAATCTCATCTCTCTTTT 1529
OY 661 caatatattgaagcattatcaggttatgtctcagagcggaatacatttgaatgt 720
DB 1528 CAATATTATTGAAGCATTTATCAGGGTTATGTCATGAGCGGATACATATTTGAATGT 1469
OY 721 attagaataaataaataaaggggttccgcgcacattcccgaaagtgcaaccgtac 780
DB 1468 ATTGAGAAAAATTAACAAATAGGGGTTCCGCGCACATTTCCCGCAAAAGTCCACCTGAC 1409
OY 781 gtctagaacacattatctcagcattacatacctataaataagcgatatacagagccc 840
DB 1408 GTCTAGAAACCATTAATATCATGACATTAACTATAAAATAGCGGATACAGAGGCC 1349
OY 841 ttctcgtcgcgcgttctcgtatatacagtgtaaaacccctacacatgcagctccggag 900
DB 1348 TTTTCGTCGCGCGCTTGGGTGATGACGCTGAAAACTCTACACATGACAGCTCCGAG 1289
OY 901 acggttcaacgcttgcgttaagcggatgcgcggagacgaacacgcgttcagggcgctca 960
DB 1288 ACGGTCAACGCTTGTCTATAGCCGATGCCCGGAGACAAACCCGTCAGCGCGCTCA 1229
OY 961 gcgggtgttgcggggtgcgcgggtgtgcttaactatgcgc 1001
DB 1228 GCGGGGTGTTGGCGGCTGTCCGGGCTGCTTACTATATGCGGC 1188

RESULT 7
AA51632/c
ID AA51632 standard; DNA: 3383 BP.
XX AC
XX AA51632;
XX DT
XX 31-OCT-2000 (first entry)
XX DE
XX Plasmid pGM678 containing E. coli deod gene.
XX KW
XX Uridine phosphorylase; uhp; purine nucleoside phosphorylase; PNP; deod.
XX KW Genetically modified organism; catalyst; transglycosylation;
XX KW nucleoside analogue; alpha-pentose-1-phosphate sugar; phosphorolysis;
XX KW anti-viral; anti-tumour; mesophilic bacterium; ss.
XX OS
XX Escherichia coli.
XX OS Synthetic.
XX FH
XX Key Location/Qualifiers
XX FT misc_RNA 1..230
XX FT /*tag= a
XX FT /label= pUC18_sequence
XX FT CDS 216..952
XX FT /*tag= b

```

```

FT      misc_RNA      /product= lacZ-deod_fused_gene
FT      961..383
FT      /tag= C
FT      /label= puc18_sequence
XX      MO200039307-A2.
XX      06-JUL-2000.
XX      23-DEC-1999; 99WO-BP10416.
XX      23-DEC-1998; 98IT-MI02792.
XX      (NORP-) NORPHARMA SPA.
XX      Bestetti G, Cali' S, Ghisotti D, Orsini G, Tonon G, Zuffi G;
XX      WPI; 2000-452402/39.
XX      Recombinant expression vectors used to transform cells for the
XX      production of e.g. nucleosides encode uridine phosphorylase and/or
XX      purine nucleoside phosphorylase
XX      Claim 14; Page 47-48; 72pp; English.
XX      Host cells genetically modified to express uridine phosphorylase (UDP)
XX      and purine nucleoside phosphorylase (PNP) or their corresponding crude or
XX      purified extracts, either separately or in combination are used as
XX      catalysts of transglycosylation reactions between a donor nucleoside and
XX      an acceptor base, for preparing nucleoside analogues containing
XX      heterocyclic systems with purine and/or pyrimidine bases substituted by
XX      one or more nitrogen atoms, for preparing alpha-pentose-1-phosphate
XX      and sugars by phosphorylase reactions and for producing nucleosides and
XX      modified nucleoside analogues (all claimed). The modified or natural
XX      nucleosides are used directly or as intermediates in the preparation of
XX      drugs with anti-viral and anti-tumour activity and for preparing
XX      oligonucleotides for therapeutic or diagnostic use. The recombinant host
XX      cells are produced using recombinant plasmid expression vectors
XX      comprising at least one gene sequence of a mesophilic bacterium coding
XX      for a polypeptide with udp activity and/or enzyme PNP activity and at
XX      least one gene sequence coding for antibiotic resistance. In particular,
XX      the E. coli udp and deod genes are used. Recombinant strains produced
XX      using the vectors express polypeptides with enzyme UDP and PNP activity
XX      in large amounts, e.g. 340-1040 times higher udp activity and 120-200
XX      times higher PNP activity than non-transformed corresponding wild type
XX      strains.
XX      Sequence 3383 BP; 826 A; 868 C; 855 G; 834 T; 0 other;
SQ

```

```

Query Match      100.0%; Score 1001; DB 21; Length 3383;
Best Local Similarity 100.0%; Pred. No. 3 8e-263;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX      301 gtcatgccatccgtaagatgtctttctgtactgtgtgactcaacaagtcacttca 360
XX      1928 gtcatgccatccgtaagatgtctttctgtactgtgtgactcaacaagtcacttca 1869
XX      361 gaatagtatgcgagcgacagagtgctcttcccgcggtcaatacgggataacacgcg 420
XX      1868 gaatagtatgcgagcgacagagtgctcttcccgcggtcaatacgggataacacgcg 1809
XX      421 ccaatagcagaactttaaaagtctcaacttggaacagttcttcggggcgaaatc 480
XX      1808 ccacatagcagaactttaaaagtctcaacttggaacagttcttcggggcgaaatc 1749
XX      481 tcaagatccttaccgctgttgagatccagtgtaacacacacgcgtgcacccaactga 540
XX      1748 tcaagatccttaccgctgttgagatccagtgtaacacacacgcgtgcacccaactga 1689
XX      541 tcttcagatccttaccgctgttgagatccagtgtaacacacacgcgtgcacccaactga 600
XX      1688 tcttcagatccttaccgctgttgagatccagtgtaacacacacgcgtgcacccaactga 1629
XX      601 gccgcaaaaaagagataagcgagacaggaatgttgaatacactactcttctt 660
XX      1628 gccgcaaaaaagagataagcgagacaggaatgttgaatacactactcttctt 1569
XX      661 caatatattgaagcatlatacagggatattgtccatgagcgaatacatattgaatgt 720
XX      1568 caatatattgaagcatlatacagggatattgtccatgagcgaatacatattgaatgt 1509
XX      721 attagaanaataaacaataaggggttcgcgcacattcccgcaaaagtcacactgtac 780
XX      1508 attagaanaataaacaataaggggttcgcgcacattcccgcaaaagtcacactgtac 1449
XX      781 gcttaagaacacattattatcatagacataacctaataatagacgtatcaagagcc 840
XX      1448 gcttaagaacacattattatcatagacataacctaataatagacgtatcaagagcc 1389
XX      841 ttctgtctgcgagttctgtgtaagcagtgtaagcagtgtaagcagtgtaagcagtgta 900
XX      1388 ttctgtctgcgagttctgtgtaagcagtgtaagcagtgtaagcagtgtaagcagtgta 1329
XX      901 acggtacagcttctgtgtaagcagtgtaagcagtgtaagcagtgtaagcagtgtaagc 960
XX      1328 acggtacagcttctgtgtaagcagtgtaagcagtgtaagcagtgtaagcagtgtaagc 1269
XX      961 gcgagtggtgcgagtggtgcgagtggtgcgagtggtgcgagtggtgcgagtggtgcgag 1001
XX      1268 gcgagtggtgcgagtggtgcgagtggtgcgagtggtgcgagtggtgcgagtggtgcgag 1228

```

```

RESULT 8
AA51630/c
ID      AA51630 standard; DNA; 3444 BP.
XX      AC      AA51630;
XX      DT      31-OCT-2000 (first entry)
XX      DE      Plasmid pGM679 containing E. coli udp gene.
XX      KW      Uridine phosphorylase; udp; purine nucleoside phosphorylase; PNP; deod;
XX      KW      Genetically modified organism; catalyst; transglycosylation;
XX      KW      nucleoside analogue; alpha-pentose-1-phosphate sugar; phosphorylase;
XX      KW      anti-viral; anti-tumour; mesophilic bacterium; ss.
XX      OS      Escherichia coli.
XX      SY      Synthetic.
XX      FH      Key
XX      FT      misc_RNA      Location/Qualifiers
XX      FT      1..242
XX      FT      /tag= a
XX      FT      /label= puc18_sequence
XX      CDS      216..1013

```

```

FT      /*tag= b
FT      /product= lacZ-udp_fused_gene
FT      1022..3444
FT      /*tag= c
FT      /label= pUC19_sequence
XX
XX      W0200039307-A2.
XX
XX      06-JUL-2000.
XX
XX      23-DEC-1999: 99WO-EP10416.
XX
XX      23-DEC-1998: 98IT-M102792.
XX
XX      (NORP-) NORPHARMA SPA.
XX
XX      Bestetti G, Calli S, Ghisotti D, Orsini G, Tonon G, Zuffi G;
XX      WPI: 2000-452402/39.
XX
XX      Recombinant expression vectors used to transform cells for the
XX      production of e.g. nucleosides encode uridine phosphorylase and/or
XX      putine nucleoside phosphorylase.
XX
XX      Claim 14: Page 43-44; 72pp; English.
XX
XX      Host cells genetically modified to express uridine phosphorylase (UDP)
XX      and putine nucleoside phosphorylase (PNP) or their corresponding crude or
XX      purified extracts, either separately or in combination are used as
XX      catalysts of transglycosylation reactions between a donor nucleoside and
XX      an acceptor base, for preparing nucleoside analogues containing
XX      heterocyclic systems with purine and/or pyrimidine bases substituted by
XX      one or more nitrogen atoms, for preparing alpha-pentose-1-phosphate
XX      sugars by phosphorylation reactions and for producing nucleosides and
XX      modified nucleoside analogues (all claimed). The modified or natural
XX      nucleosides are used directly or as intermediates in the preparation of
XX      drugs with anti-viral and anti-tumour activity and for preparing
XX      oligonucleotides for therapeutic or diagnostic use. The recombinant host
XX      cells are produced using recombinant plasmid expression vectors
XX      comprising at least one gene sequence of a mesophilic bacterium coding
XX      for a polypeptide with UDP activity and/or enzyme PNP activity and at
XX      least one gene sequence coding for antibiotic resistance. In particular,
XX      the E. coli udp and deoD genes are used. Recombinant strains produced
XX      using the vectors express polypeptides with enzyme UDP and PNP activity
XX      in large amounts, e.g. 340-1040 times higher UDP activity and 120-200
XX      times higher PNP activity than non-transformed corresponding wild type
XX      strains.
XX
XX      Sequence 3444 BP; 820 A; 881 C; 887 G; 856 T; 0 other:

```

```

Query Match      100.0%; Score 1001; DB 21; Length 3444;
Best Local Similarity 100.0%; Pred. No. 3.8e-263;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY      1 cctgcaacttaccgctccatccagcttattatgttgcgggaagctgagtaagt 60
DB      2289 CCTGCAACTTTATCCGCTCCATCCAGTCTATTATTTGCGGGAACTGAGTAAGT 2230
OY      61 agtcgcagtaataagttgtgcgaagctgtgtccatgtctacagagcctgtgtgta 120
DB      2229 AGTTGCCAGATTAAATAGTTGGGCAAGCTTGTTGCCATTGCTACAGGCAATCGTGGTGTCA 2170
OY      121 cgtctgtgtgtgtatgagcttcatcagctcogtltcccaagaatcaagcgagttaca 180
DB      2169 CGCTCGTCTGTTGGTATGCGATCATTCACGCTCCGTTCCCAAGATCAAGGCGAGTTACA 2110
OY      181 tgaatcccatgttgcgaataaagaggttaagctccttcgtgtccctccagatgtgtcaga 240
DB      2109 TGAATCCCATGTTGTGCAAAAAGGAGTTAGCTCTTCGTGCTCCGATCTGTGTGACA 2050
OY      241 agtaagttgcgcagtgatcatcactcagtgatagcgagcctcataactcttact 300

```

```

DB      2049 AGTAAGTGGCCGACAGTGTATCTACTCATGTTATGAGCAGACAGTCGATTAATTTCTTACT 1990
OY      301 gttatgccatccgtaagaagcttttctgtactgtgtgaactcaaccaagtcattcga 360
DB      1989 GTGATGCCATCCGTTAGAGAGCTTTTCTGTGACTGTGATGATCTCAACCAAGCATTTCTGA 1930
OY      361 gaatagtatgcgagcagagttgtctcttgcggcgctcaatacggatataacgcg 420
DB      1929 GAAATAGTATGCGGCGAGCAGGAGTCTTTCGCCGCGCTCAATACGGGATATATACCCGC 1870
OY      421 ccacatagcagaacatttaaaagtgcatacaltggaataacgttcttcggygcgaatactc 480
DB      1869 CCACATAGCAGAACTTTAAAGTGCATCATTTGGAATAACGTTCTTCGGGGGGAATACTC 1810
OY      481 tcaagatcttaacgctgtgtgagatccagttgtagtaaccactcgtgcacccaactga 540
DB      1809 TCAAGATCTTACCGCTGTGAGATCCAGTTGAGTAACTCCACTGTGACCCCAACTGA 1750
OY      541 tcttcagcatcttacttaccagcgtttctcgggtgagcaaaaacaggaagcaaat 600
DB      1749 TCTTGAGCATCTTTTACTTTACACAGCGCTTCTGGGTAGCAAAAACAGAAAGCAAAAT 1690
OY      601 gccgcataaaaggaataagggcgacaggaatgttgaactatactcttcccttt 660
DB      1689 GCCGCAAAAAGGAATAAGGCGACAGGAAATGTAATCATCTCTCTCTTTT 1630
OY      661 caatatattgaagattatcatcaggtattgtctcagcagatatactattgaatgt 720
DB      1629 CAATATTATTGAAGATTATTCAGGGTTATGTCTCATGAGCGGATACATATTGGAATGT 1570
OY      721 attagaataataacaataaggggttcgcgcacattcccgcaaaaagtgccactgac 780
DB      1569 ATTTGAAAATAATAACAATAGGGTTCGCGCACATTTCCCGAAAAGTGCCACACTGAC 1510
OY      781 gcttaagaacattatatacatatgaactataaataagtcgtatacagaagccc 840
DB      1509 GTCTAAGAAACCATTTATATGATGACATTAACCTATAAAATAGCGGTATCACAGGCC 1450
OY      841 ttctcgcgcgcttctgtgtatgaagtgtaaaacctctacacatgacgtccggag 900
DB      1449 TTTGCTTCGCGCGCTTTGCGTATGATGACGTTAATAACCTCTACACATGACCTCCGGAG 1390
OY      901 acggtcacagcttctgttgaagcgaatgcggagcagagaacccgttcagggcgctca 960
DB      1389 ACGGTCAACAGCTTGTCTCTTAAGCGGATGCCGGAGACAAAGCCGTCAGGCGCGTCA 1330
OY      961 gcgggtgttgcgggtgtcgcggggtgtgcttaactaactgagcgcg 1001
DB      1329 GCGGCTGTGGCGGTGTGCGGGCTGGCTTAATATGCGGC 1289

```

```

RESULT 9
AAT78825
ID      AAT78825 standard; DNA; 3819 BP.
AC      AAT78825;
XX
XX      23-JAN-1998 (first entry)
XX
XX      Kappa light chain plasmid pUC65.
XX
XX      Ig: affinity constant; human; antigen; hybridoma; B cell; transgene;
XX      transgenic; mouse; CD4; antibody; autoimmune; inflammatory;
XX      transplant rejection; immunoglobulin; ss.
XX      Synthetic.
XX      Homo sapiens.
XX      W09713852-A1.
XX      17-Apr-1997.
XX      10-Oct-1996; 96WO-US16433.

```



XX 10-OCT-1995; 950S-0544404.  
 XX (GENP-) GENPHARM INT INC.  
 XX Kay RM, Lonberg N;  
 XX WPI; 1997-235888/21.  
 XX Novel anti-CD4 antibody produced by transgenic mice - used in the  
 PT treatment of auto-immune disease etc.  
 XX  
 XX Example 42; Page 266-268; 396pp; English.  
 XX  
 XX A novel composition has been developed which comprises an immunoglobulin  
 CC (Ig) having an affinity constant (Ka) of at least 2 multiply  
 CC 1000000000 M-1 for binding to a predetermined human antigen. The  
 CC present sequence represents the kappa light chain plasmid pLC665  
 CC which includes the kappa constant region and polyadenylation site. Anti-  
 CC CD4 antibodies may be used in therapeutic and diagnostic applications,  
 CC especially for the treatment of human diseases. These antibodies reduce  
 CC activity of CD4 cells and reduce undesirable autoimmune reactions,  
 CC inflammatory response and transplant rejection. Transgenic animals are  
 CC capable of producing heterologous antibodies of multiple isotypes by  
 CC undergoing isotype switching. These animals produce a first Ig type  
 CC that is necessary for antigen-stimulated B-cell maturation and can  
 CC switch to encode and produce one or more subsequent heterologous  
 CC isotypes.  
 XX  
 SO Sequence 3819 BP; 947 A; 1015 C; 912 G; 945 T; 0 other;

Query Match 100.0%; Score 1001; DB 18; Length 3819;  
 Best Local Similarity 100.0%; Pred. No. 3.9e-263;  
 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cctgcgaacttaccgctccatccagctcattatgttgcgggaagctagtagtaagt 60  
 DB 1156 cctgcgaacttaccgctccatccagctcattatgttgcgggaagctagtagtaagt 1215

QY 61 agtgcgaacttaccgctccatccagctcattatgttgcgggaagctagtagtaagt 120  
 DB 1216 agtgcgaacttaccgctccatccagctcattatgttgcgggaagctagtagtaagt 1275

QY 121 cgtgcgtcgttgcgtatgcgttcattacgctgcgttcacgaagcagaagcagaagcaga 180  
 DB 1276 cgtgcgtcgttgcgtatgcgttcattacgctgcgttcacgaagcagaagcagaagcaga 1335

QY 181 tgatcccccattgtgtgcaaaaagcgtgtagctccttcgctccgcagatcgttgcaga 240  
 DB 1336 tgatcccccattgtgtgcaaaaagcgtgtagctccttcgctccgcagatcgttgcaga 1395

QY 241 agtaagtgtgcgcagtggttacaactcattatgttgcgggaagctagtagtaagt 300  
 DB 1396 agtaagtgtgcgcagtggttacaactcattatgttgcgggaagctagtagtaagt 1455

QY 301 gtcagtcacccggaagatccttcttgcgtacgtgtagtaccacgaagcagaagcaga 360  
 DB 1456 gtcagtcacccggaagatccttcttgcgtacgtgtagtaccacgaagcagaagcaga 1515

QY 361 gaatagtgatgcgcgcagcagtggttgcgttccttcgctccgcagcagaagcaga 420  
 DB 1516 gaatagtgatgcgcgcagcagtggttgcgttccttcgctccgcagcagaagcaga 1575

QY 421 ccaatagcagaacttcaaaaagtgctcattatgttgcgggaagctagtagtaagt 480  
 DB 1576 ccaatagcagaacttcaaaaagtgctcattatgttgcgggaagctagtagtaagt 1635

QY 481 tcaaggaacttaccgctgttgcgtacgttgcgtacgttgcgtacgttgcgtacgttgcgt 540  
 DB 1636 tcaaggaacttaccgctgttgcgtacgttgcgtacgttgcgtacgttgcgtacgttgcgt 1695

QY 541 tcttgaagcatttacttaccagcgttcttcgtgtgagcaaaaagcagaagcagaat 600

DB 1696 tcttgaagcatttacttaccagcgttcttcgtgtgagcaaaaagcagaagcagaat 1755  
 QY 601 gccgcaaaaagcagaagcagaagcagaagcagaagcagaagcagaagcagaagcaga 660  
 DB 1756 gccgcaaaaagcagaagcagaagcagaagcagaagcagaagcagaagcagaagcaga 1815  
 QY 661 caatattatgaagcattatcagcgttattgtctcgtgcgtgcgtgcgtgcgtgcgt 720  
 DB 1816 caatattatgaagcattatcagcgttattgtctcgtgcgtgcgtgcgtgcgtgcgt 1875  
 QY 721 attagaanaatacaaatagcgttcgcgcagcattccgcggaagcagcagcagc 780  
 DB 1876 attagaanaatacaaatagcgttcgcgcagcattccgcggaagcagcagcagc 1935  
 QY 781 gcttaagaacacattatcattacattacattacattacattacattacattacattac 840  
 DB 1936 gcttaagaacacattatcattacattacattacattacattacattacattacattac 1995  
 QY 841 ttgcgtcgcgcgttcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgt 900  
 DB 1996 ttgcgtcgcgcgttcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgt 2055  
 QY 901 acggtcacagcttgcgttcgttcgttcgttcgttcgttcgttcgttcgttcgttcgt 960  
 DB 2056 acggtcacagcttgcgttcgttcgttcgttcgttcgttcgttcgttcgttcgttcgt 2115  
 QY 961 gcgggtgttgcgcgttcgttcgttcgttcgttcgttcgttcgttcgttcgttcgt 1001  
 DB 2116 gcgggtgttgcgcgttcgttcgttcgttcgttcgttcgttcgttcgttcgttcgt 2156

RESULT 10  
 AAV39266  
 ID AAV39266 standard; DNA: 3819 BP.  
 XX AAV39266;  
 XX 18-DEC-1998 (first entry)  
 XX  
 DE Plasmid pLC665 nucleotide sequence.  
 XX  
 OS Transgenic animal; human heterologous antibody; transgene;  
 XX isotype switching; neutrophil efflux; reperfusion injury; CD4 binding;  
 KW autoimmune reaction; inflammatory response; transplant rejection;  
 KW acid induced lung injury; acute adult respiratory distress syndrome;  
 KW ARDS; vasculitis; septic shock; allergic reaction; asthma;  
 KW cystic fibrosis; ss.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 XX  
 PN MO9824B84-A1.  
 PD 11-JUN-1998.  
 XX  
 PF 01-DEC-1997; 97WO-US21803.  
 XX  
 PR 02-DEC-1996; 96US-0758417.  
 XX  
 PA (GENP-) GENPHARM INT.  
 XX  
 PI Kay RM, Lonberg N;  
 XX  
 DR WPI; 1998-333306/29.  
 XX  
 PT Hybridoma producing antibody specific for interleukin-8 - used to  
 PT prevent efflux of neutrophils from vasculature, and treat  
 PT reperfusion injury  
 XX  
 PS Example 42; Pages 317-319; 452pp; English.  
 XX  
 CC The present sequence represents a plasmid, pLC665, which contains a

CC synthetic kappa light chain sequence (created using oligonucleotide  
 CC AAV39244-65). This synthetic sequence differs from natural sequences in  
 CC that strings of repeated oligonucleotides are interrupted (to facilitate  
 CC oligonucleotide synthesis and PCR amplification), optimal translation  
 CC initiation sites are incorporated and HindIII sites were engineered  
 CC upstream of the translation initiation sites. The plasmid is used in the  
 CC construction of minigenes for expression of Igkappa anti-CD4  
 CC antibodies, in the transgenic mouse of the invention. The specification  
 CC describes transgenic non-human animals, especially a mouse, which are  
 CC capable of producing a human heterologous antibodies of multiple isotypes  
 CC by undergoing isotype switching. The transgenic animals have human heavy  
 CC and light chain transgenes. The transgenes are capable of functionally  
 CC rearranging a heterologous diversity (D) gene in a  
 CC variable-diversity-junction (V-D-J) recombination. The transgenes  
 CC include a heavy chain transgene comprising at least one V, D and J gene  
 CC segment, and one constant region gene segment. The immunoglobulin (Ig)  
 CC light chain transgene comprises at least one V and J gene segment and one  
 CC constant region gene segment. The gene segments are heterologous to the  
 CC transgenic animal. The antibody can be used to prevent efflux of  
 CC neutrophils from vasculature. It can also be used to treat reperfusion  
 CC injury. CD4 binding antibodies are used to reduce undesirable autoimmune  
 CC reactions, inflammatory responses and rejection of transplantable organs.  
 CC The anti-IL-8 antibodies can reduce tissue damage and prolong survival  
 CC in animal models of acute adult respiratory distress syndrome (ARDS) and  
 CC acid induced lung injury. The anti-IL-8 antibodies can also be used for  
 CC the treatment of cystic fibrosis, septic shock, allergic reactions  
 CC (e.g. asthma) and vascular fibrosis.

XX Sequence 3819 BP; 947 A; 1015 C; 912 G; 945 T; 0 other:

Query Match 100.0%; Score 1001; DB 19; Length 3819;

Best Local Similarity 100.0%; Pred. No. 3.9e-263; Mismatches 0; Indels 0; Gaps 0;

Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ccgcgaacttaccgctccatccagctcattatgttgcgggaagctagaatg 60  
 DB 1156 ccgcgaacttaccgctccatccagctcattatgttgcgggaagctagaatg 1215

QY 61 agtgcgcagtaataatgttgcgaacgtgtgtgcattgtcacaagcactcggtgtca 120  
 DB 1216 agtgcgcagtaataatgttgcgaacgtgtgtgcattgtcacaagcactcggtgtca 1275

QY 121 cgtctcgtttgttgttgttattcagctcagctcgcgtcccaagatacgaagcagttaca 180  
 DB 1276 cgtctcgtttgttgttgttattcagctcagctcgcgtcccaagatacgaagcagttaca 1335

QY 181 tgatcccccattgttgcgaacagcgttagctcctcgtctcctcagtcgtgtcaga 240  
 DB 1336 tgatcccccattgttgcgaacagcgttagctcctcgtctcctcagtcgtgtcaga 1395

QY 241 agtaagtgtgcgcagctgttatacctcaatggtatgagcagcactgatactcttact 300  
 DB 1396 agtaagtgtgcgcagctgttatacctcaatggtatgagcagcactgatactcttact 1455

QY 301 gtcattgcacccgttaagatgacttcttgcactggtgagttactcaacacgacttttga 360  
 DB 1456 gtcattgcacccgttaagatgacttcttgcactggtgagttactcaacacgacttttga 1515

QY 361 gaatagtgtatgcgcagcagctgtctcttgcgcgcgtcacaatacggatataccggc 420  
 DB 1516 gaatagtgtatgcgcagcagctgtctcttgcgcgcgtcacaatacggatataccggc 1575

QY 421 ccacataagcagaacttaaaagtgtcattcatcttggaacacgttcttcgggcgaacac 480  
 DB 1576 ccacataagcagaacttaaaagtgtcattcatcttggaacacgttcttcgggcgaacac 1635

QY 481 tcaagatcttaccgctgttgagatccagttgagatgtaaccacccgtgcaccacacga 540  
 DB 1636 tcaagatcttaccgctgttgagatccagttgagatgtaaccacccgtgcaccacacga 1695

QY 541 tcttcagcatcttacttccacagcgttcttggtgagcaaaaacggaagcaaat 600  
 DB 541 tcttcagcatcttacttccacagcgttcttggtgagcaaaaacggaagcaaat 600

DB 1696 tcttcagcatcttacttccacagcgttcttggtgagcaaaaacggaagcaaat 1755

QY 601 gccgcaaaaaaggaataaggcgacacggaatgttgatacctaactcttctt 660

DB 1756 gccgcaaaaaaggaataaggcgacacggaatgttgatacctaactcttctt 1815

QY 661 caatatattgaagcattatcaagggttatgttccatgaagcggatacatttgatgt 720

DB 1816 caatatattgaagcattatcaagggttatgttccatgaagcggatacatttgatgt 1875

QY 721 attagaataataacaaataagggttccgcgcacattcccccgaagaagtgcacctgac 780

DB 1876 attagaataataacaaataagggttccgcgcacattcccccgaagaagtgcacctgac 1935

QY 781 gcttaagaacacattatctcagacattacctaataaataaggcgtatcaagggccc 840

DB 1936 gcttaagaacacattatctcagacattacctaataaataaggcgtatcaagggccc 1995

QY 841 ttctcgtcgcgcgttctcgttgatgacggtgaaacccctgcacacatgcactccggag 900

DB 1996 ttctcgtcgcgcgttctcgttgatgacggtgaaacccctgcacacatgcactccggag 2055

QY 901 acggtcacagcttctcgttaagcggatgcgggagcagaacccgtcaggcgctga 960

DB 2056 acggtcacagcttctcgttaagcggatgcgggagcagaacccgtcaggcgctga 2115

QY 961 gcgggtgttgccgggtgttcggagcgtgacttaactatgcgc 1001

DB 2116 gcgggtgttgccgggtgttcggagcgtgacttaactatgcgc 2156

RESULT 11

AAZ22020

ID AAZ22020 standard; DNA: 3819 BP.

XX AAZ22020;

DE 24-NOV-1999 (first entry)

XX Nucleotide sequence of plasmid pUC65.

KW Transgenic animal; heterologous antibody; hybridoma; B cell;

KW transgenic mouse; human heavy chain transgene; digoxin;

KW human light chain transgene; immortalized cell; immunoglobulin;

KW Shiga-like toxin; autoimmune disease; cancer; infectious disease;

KW transplant rejection; blood disorder; coagulation disorder; ss.

OS Synthetic.

XX W09945962-A1.

XX 16-SEP-1999.

XX 12-MAR-1999; 99MO-US0535.

XX 13-MAR-1998; 98US-0042353.

XX (GENP-) GENPHARM INT. INC.

XX Lonberg N, Fishwild DM, Ball WT;

XX WPI: 1999-551219/46.

XX Novel transgenic non-human animals used to produce heterologous

XX antibodies

XX Example 42; Page 318-320; 484pp; English.

XX The specification describes transgenic animals that are capable of

XX producing a heterologous antibody. The antibodies are isolated from a

XX hybridoma, comprising B cells, that is obtained from a transgenic mouse

XX having a genome comprising a human heavy chain transgene and a human

XX light chain transgene. The B cells are fused to immortalized cells

CC suitable for generating a hybridoma, which produces a detectable  
CC amount of an immunoglobulin that specifically binds dioxin or  
CC Shiga-like toxin. B cells from transgenic animals can be used to  
CC generate hybridomas expressing monoclonal high affinity human sequence  
CC antibodies. Antibodies produced from the transgenic animals of the  
CC invention can be used to treat human diseases, e.g. autoimmune  
CC diseases, cancer, infectious disease, transplant rejection, blood  
CC disorders such as coagulation disorders and other diseases. The  
CC present sequence is used in the course of the invention.

XX Sequence 3819 BP; 947 A; 1015 C; 912 G; 945 T; 0 other;

Query Match 100.0%; Score 1001; DB 20; Length 3819;  
Best Local Similarity 100.0%; Pred. No. 3.9e-263;  
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cctgcacattatccgcctccatccagctctatattgttgcgggaagtagaagt 60  
DB 1156 cctgcacattatccgcctccatccagctctatattgttgcgggaagtagaagt 1215  
QY 61 agtgcacgttaatagtttgcgcaactgtttgcattgtctacagcgatcgtgttca 120  
DB 1216 agtgcacgttaatagtttgcgcaactgtttgcattgtctacagcgatcgtgttca 1275  
QY 121 cgtcgtcgtttgtatagcttcatcagctccggtcccaacgatacagcgagttaca 180  
DB 1276 cgtcgtcgtttgtatagcttcatcagctccggtcccaacgatacagcgagttaca 1335  
QY 181 tgatcccccattgttgcacaaaagcggttagctccttcgtctccgcgtgttcaga 240  
DB 1336 tgatcccccattgttgcacaaaagcggttagctccttcgtctccgcgtgttcaga 1395  
QY 241 agtaagtgtgcgagtggtatcacctcagtggtatggcagcctgcataattcttact 300  
DB 1396 agtaagtgtgcgagtggtatcacctcagtggtatggcagcctgcataattcttact 1455  
QY 301 gtcattccatccgtaagaagctttctgctgactgtgtaactcaaccaagtcattctga 360  
DB 1456 gtcattccatccgtaagaagctttctgctgactgtgtaactcaaccaagtcattctga 1515  
QY 361 gaatagtgtatgcgcgcagagttgctcttgcgcgcgtcaatacggagataacgcgcg 420  
DB 1516 gaatagtgtatgcgcgcagagttgctcttgcgcgcgtcaatacggagataacgcgcg 1575  
QY 421 ccacatagcagaactttaaagtgtcctacatttgaaaaagctttctcggcggaanaactc 480  
DB 1576 ccacatagcagaactttaaagtgtcctacatttgaaaaagctttctcggcggaanaactc 1635  
QY 481 tcaagatcttaacgctgttgagatcagatcgatgtaaccacactcgttcaccacaactga 540  
DB 1636 tcaagatcttaacgctgttgagatcagatcgatgtaaccacactcgttcaccacaactga 1695  
QY 541 tcttcagcatcttacttaccacagcgtttctgggtgagcaaaaaacaggaagcaacaa 600  
DB 1696 tcttcagcatcttacttaccacagcgtttctgggtgagcaaaaaacaggaagcaacaa 1755  
QY 601 gcgcgcaaaaaggaataaggcgacaggaatgttgatactctatcttctt 660  
DB 1756 gcgcgcaaaaaggaataaggcgacaggaatgttgatactctatcttctt 1815  
QY 661 caatatatgaagcattatcagagttatgtctcatagcagatatacttgaatgt 720  
DB 1816 caatatatgaagcattatcagagttatgtctcatagcagatatacttgaatgt 1875  
QY 721 atttagaaaaataacaaaagggttcgcgcgacattcccgaaaagtgcacactgac 780  
DB 1876 atttagaaaaataacaaaagggttcgcgcgacattcccgaaaagtgcacactgac 1935  
QY 781 gtctaaagaacatttatctgacatataactataaaatagggatcagagagccc 840  
DB 1936 gtctaaagaacatttatctgacatataactataaaatagggatcagagagccc 1995

QY 841 ttctgtcgcgcgcttctcgtatgacggtgaaaaactctgacacatgcagctccgcgag 900  
DB 1996 ttctgtcgcgcgcttctcgtatgacggtgaaaaactctgacacatgcagctccgcgag 2055  
QY 901 acggtcacagcttctcgtatgacggtgacggagacagacagccgttcaggcgctca 960  
DB 2056 acggtcacagcttctcgtatgacggtgacggagacagacagccgttcaggcgctca 2115  
QY 961 gcgggtgttgcggggttgcggagctggtgcttaactatgcgc 1001  
DB 2116 gcgggtgttgcggggttgcggagctggtgcttaactatgcgc 2156

# RESULT 12

AAD04947/c  
ID AAD04947 standard; DNA; 3858 BP.

AAD04947;

17-JUL-2001 (first entry)

DE Plasmid pRK50 used to test Cre recombinase mediated inversion.

XX Gene trapping construct: conditional mutation; unidirectional inversion;  
XX recombinease recognition sequence; RRS; disruption cassette;  
XX selection cassette; transgenic organism; pRK50; Cre recombinase; ds.

OS Chimeric - Adenovirus.

OS Chimeric - Unidentified.

OS Chimeric - ECMV virus.

PN W0200129208-A1.

PD 26-APR-2001.

XX 16-OCT-2000; 2000WO-EP10162.

XX 16-OCT-1999; 99EP-0120592.

PR 27-OCT-1999; 99US-0162016.

PA (ARTE-) ARTEMIS PHARM GMBH.

FR (FRAN-) FRANKGEN BIOTECHNOLOGIE AG.

PI Kuehn R, Von Melchner H, Altschmied J;

XX WPI: 2001-308486/32.

XX New gene trapping construct capable of causing conditional mutations in

XX genes, comprises functional DNA segment inserted in sense or antisense

XX direction relative to gene to be trapped

XX Example 3; Page 73-74; 78pp; English.

CC The present invention relates to a conditional gene trapping construct  
CC capable of causing conditional mutations in genes. The gene trapping  
CC construct comprises two functional DNA segments, each being flanked by  
CC two recombinase recognition sequences (RRS) specific to site specific  
CC recombinase which is capable of unidirectional inversion of double  
CC standard DNA segment. One of the DNA segment (disruption cassette) is  
CC inserted in antisense orientation relative to the transcriptional  
CC orientation of the gene to be trapped. The other DNA segment (selection  
CC cassette) is inserted in sense direction relative to the transcriptional  
CC orientation of the gene to be trapped. The cell comprising the gene  
CC trapping construct is useful for the identification and/or isolation of  
CC genes. The transgenic organism comprising the gene trapping construct is  
CC useful to study gene function at various developmental stages. The gene  
CC trapping construct is useful for mutationally inactivating all cellular  
CC genes. The present sequence is pRK50 vector, which is used to test Cre  
CC recombinase mediated inversion, which is related to the invention.

XX Sequence 3858 BP; 943 A; 959 C; 994 G; 962 T; 0 other;

Query Match 100.0%; Score 1001; DB 22; Length 3858;  
 Best Local Similarity 100.0%; Pred. No. 3,9e-263;  
 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 cctgaacttaccgcccgcacacgctctatattgttcgagggaagcagaagtaagt 60
DB 2435 cctgaacttaccgcccgcacacgctctatattgttcgagggaagcagaagtaagt 2376
QY 61 agttcgcaagtaataagtttcgcaaggtttgttcacattgttcacagcatcgttgttca 120
DB 2375 agttcgcaagtaataagtttcgcaaggtttgttcacattgttcacagcatcgttgttca 2316
QY 121 cgtctcgtttgtatagttcattcaagtcgcgttcccaagacataaggcaggttaca 180
DB 2315 cgtctcgtttgtatagttcattcaagtcgcgttcccaagacataaggcaggttaca 2256
QY 181 tgatcccccatgtttgtgtaaaaaagggtttgctcccttggttcctccatcgttgtaga 240
DB 2255 tgatcccccatgtttgtgtaaaaaagggtttgctcccttggttcctccatcgttgtaga 2196
QY 241 agtaagttgcccagagtttatacactcatgtttatgcaagcactgcatattctctact 300
DB 2195 agtaagttgcccagagtttatacactcatgtttatgcaagcactgcatattctctact 2136
QY 301 gtcacgtccatcgttaagaatgctttctgtgacgtgtgagttacatcaacaaagttcgtga 360
DB 2135 gtcacgtccatcgttaagaatgctttctgtgacgtgtgagttacatcaacaaagttcgtga 2076
QY 361 gaatagtatgtagcgagcagcaggttgccttgcgcgcgtcaataaggagataaccgcg 420
DB 2075 gaatagtatgtagcgagcagcaggttgccttgcgcgcgtcaataaggagataaccgcg 2016
QY 421 ccacatagcagaacttttaaaagtctacatcatcttgtaaaacgttcttcgagcgaaacatc 480
DB 2015 ccacatagcagaacttttaaaagtctacatcatcttgtaaaacgttcttcgagcgaaacatc 1956
QY 481 tcaagatcttaacgcgtgttgagatccaagttcgatgtaaccacatcgtgcaaccaactga 540
DB 1955 tcaagatcttaacgcgtgttgagatccaagttcgatgtaaccacatcgtgcaaccaactga 1896
QY 541 tcttcagaacatcttacttaccacagcgtttctgtgtgagcaaaaaaggaaagcaaat 600
DB 1895 tcttcagaacatcttacttaccacagcgtttctgtgtgagcaaaaaaggaaagcaaat 1836
QY 601 gccgcaaaaaaggaaataaaggcgacacaggaatgttgaatcactcactcttcttctt 660
DB 1835 gccgcaaaaaaggaaataaaggcgacacaggaatgttgaatcactcactcttcttctt 1776
QY 661 caatataatgaagcattatcaaggttatgtctcatgagcgagatacatatttgaatgt 720
DB 1775 caatataatgaagcattatcaaggttatgtctcatgagcgagatacatatttgaatgt 1716
QY 721 atttagaaaaataaagaatagggttccgcgcacatttccccaagaagtgccacatgac 780
DB 1715 atttagaaaaataaagaatagggttccgcgcacatttccccaagaagtgccacatgac 1656
QY 781 gttcaagaanaacatattatcatgacatlaaccataaaaaaagaggtatcaacagagccc 840
DB 1655 gttcaagaanaacatattatcatgacatlaaccataaaaaaagaggtatcaacagagccc 1596
QY 841 ttctcgtcgcggttccgtgtgacaggtgaaaaacctgtgacacatgcagctccggag 900
DB 1595 ttctcgtcgcggttccgtgtgacaggtgaaaaacctgtgacacatgcagctccggag 1536
QY 901 acggtcacagctgtctgttaagcgagatgcgagagcagacaagccgtcaaggcggttca 960
DB 1535 acggtcacagctgtctgttaagcgagatgcgagagcagacaagccgtcaaggcggttca 1476
QY 961 gcggagttgtgcggggtcgtgcggggttgcggttcaacatgcagc 1001
DB 1475 gcggagttgtgcggggtcgtgcggggttgcggttcaacatgcagc 1435

```

RESULT 13  
 AAV39242  
 ID AAV39242 standard; DNA; 3880 BP.  
 XX  
 XX

AAV39242:

18-DEC-1998 (first entry)

Plasmid pCK7-96 nucleotide sequence.

Transgenic animal; human heterologous antibody; transgene:  
 isotype switching; neutrophil efflux; reperfusion injury; CD4 binding;  
 autoimmune reaction; inflammatory response; transplant rejection;  
 acid induced lung injury; acute adult respiratory distress syndrome;  
 ARDS; vasculitis; septic shock; allergic reaction; asthma;  
 cystic fibrosis; ss.

Synthetic.

Homo sapiens.

MO824884-A1.

11-JUN-1998.

01-DEC-1997; 97WO-US21803.

02-DEC-1996; 96US-0758417.

(GENP-) GENPHARM INT.

Kay RM, Lonberg N;

WPI; 1998-333306/29.

Hybridoma producing antibody specific for Interleukin-8 - used to  
 prevent efflux of neutrophils from vasculature, and treat  
 reperfusion injury

Example 42; Pages 310-312; 452pp; English.

The present sequence represents the kappa light chain plasmid, pCK7-96,  
 which includes the kappa constant region and polydenylation site. The  
 plasmid is used in the construction of minigenes for expression of  
 Igkappa anti-CD4 antibodies, in the transgenic mouse of the invention.  
 The specification describes transgenic non-human animals, especially a  
 mouse, which are capable of producing a human heterologous antibodies of  
 multiple isotypes by undergoing isotype switching. The transgenic animals  
 have human heavy and light chain transgenes. The transgenes are capable  
 of functionally rearranging a heterologous diversity (D) gene in a  
 variable-diversity-junction (V-D-J) recombination. The transgenes include  
 a heavy chain transgene comprising at least one V, D and J gene segment,  
 and one constant region gene segment. The immunoglobulin (Ig) light chain  
 transgene comprises at least one V and J gene segment and one constant  
 region gene segment. The gene segments are heterologous to the transgenic  
 animal. The antibody can be used to prevent efflux of neutrophils from  
 vasculature. It can also be used to treat reperfusion injury. CD4 binding  
 antibodies are used to reduce undesirable autoimmune reactions,  
 inflammatory responses and rejection of transplanted organs. The  
 anti-IL-8 antibodies can reduce tissue damage and prolong survival in  
 animal models of acute adult respiratory distress syndrome (ARDS) and  
 acid induced lung injury. The anti-IL-8 antibodies can also be used for  
 the treatment of vasculitis, septic shock, allergic reactions  
 (e.g. asthma) and cystic fibrosis.

Sequence 3880 BP; 981 A; 1013 C; 921 G; 965 T; 0 other;

Query Match 100.0%; Score 1001; DB 19; Length 3880;

Best Local Similarity 100.0%; Pred. No. 4e-263;

Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 cctgaacttaccgcccgcacacgctctatattgttcgagggaagcagaagtaagt 60
  |||

```

```

Db 1156 cctgcgaacttatccgcctccatccagctctatattgtgtgcgggaagctagatgaat 1215
QY 61 agtctgcagatlaatatgtttgcgaacgttgttgcattgtctaacaggaatcgttgtgtca 120
Db 1216 agtctgcagatlaatatgtttgcgaacgttgttgcattgtctaacaggaatcgttgtgtca 1275
QY 121 cgtctcgttgttgtatgtcttcatcagctcgcgtgtcccaagaatcagaagcgagttaaa 180
Db 1276 cgtctcgttgttgtatgtcttcatcagctcgcgtgtcccaagaatcagaagcgagttaaa 1335
QY 181 tgaatccccatgttgtgcaaaaagcggttagctctcgtctcgtctcgtctcgtctcgtcaga 240
Db 1336 tgaatccccatgttgtgcaaaaagcggttagctctcgtctcgtctcgtctcgtctcgtcaga 1395
QY 241 agtaagttgcccagctgttatcaactaatgttatacgttcgagcaactgataatcttact 300
Db 1396 agtaagttgcccagctgttatcaactaatgttatacgttcgagcaactgataatcttact 1455
QY 301 gtcatagcacgttaagatgtcttctgtgactggtgagtaactcaacacgaatctctga 360
Db 1456 gtcatagcacgttaagatgtcttctgtgactggtgagtaactcaacacgaatctctga 1515
QY 361 gaatagttatacgttcgagcaactgtctcttgcggcggtcaatacaggaataacgcgcg 420
Db 1516 gaatagttatacgttcgagcaactgtctcttgcggcggtcaatacaggaataacgcgcg 1575
QY 421 ccacatagcagaacttaaaagtggtctcatctatgtgaaaacgtttcttcggggcgaaaactc 480
Db 1576 ccacatagcagaacttaaaagtggtctcatctatgtgaaaacgtttcttcggggcgaaaactc 1635
QY 481 tcaagagcttaccgctggttgagatccagctgcatgtgaacccactcgtgacccaactga 540
Db 1636 tcaagagcttaccgctggttgagatccagctgcatgtgaacccactcgtgacccaactga 1695
QY 541 tcttcagcatcttacttaccacagcggtttctggtgtgagcaaaaacaggaagggcaaat 600
Db 1696 tcttcagcatcttacttaccacagcggtttctggtgtgagcaaaaacaggaagggcaaat 1755
QY 601 gccgcgaaaaaggaataaaggcgacacggaatgttgaatactactctctctt 660
Db 1756 gccgcgaaaaaggaataaaggcgacacggaatgttgaatactactctctctt 1815
QY 661 caaatattgaagacattatccaggttactgtctcatcagcgagatacatatttgaatgt 720
Db 1816 caaatattgaagacattatccaggttactgtctcatcagcgagatacatatttgaatgt 1875
QY 721 attagaataataaacaataaggggttcgcgcgaacattcccccgaagaatgtgcacactgac 780
Db 1876 attagaataataaacaataaggggttcgcgcgaacattcccccgaagaatgtgcacactgac 1935
QY 781 gtctaaagaacattatcatcacaactaaactataaataaggggttacaagaagccc 840
Db 1936 gtctaaagaacattatcatcacaactaaactataaataaggggttacaagaagccc 1995
QY 841 ttctcgtcgcgcgtttccgttgtatgacgtgtgaaaacctctgacacatgacgtcccgag 900
Db 1996 ttctcgtcgcgcgtttccgttgtatgacgtgtgaaaacctctgacacatgacgtcccgag 2055
QY 901 acggtcacacgttgtctgtatgacgtgtgacggagacagaacacccgttcagggcggtca 960
Db 2056 acggtcacacgttgtctgtatgacgtgtgacggagacagaacacccgttcagggcggtca 2115
QY 961 gcgggtgttgcgggtgtcggggcgtgcttaactatcgccg 1001
Db 2116 gcgggtgttgcgggtgtcggggcgtgcttaactatcgccg 2156

```

## RESULT 14

AAT78801

ID AAT78801 standard; DNA: 3881 BP.

XX AAT78801;

XX

```

DF 23-JAN-1998 (first entry)
DE Kappa light chain plasmid pCK7-96.
XX Ig; affinity constant; human; antigen; hybridoma; B cell; transgene;
XX transgenic; mouse; CD4; antibody; autoimmunity; inflammatory;
KW transplant rejection; immunoglobulin; ss.
XX Synthetic.
OS Homo sapiens.
XX MO9713852-A1.
XX 17-APR-1997.
XX 10-OCT-1996; 96WO-US16433.
XX 10-OCT-1995; 95US-0544404.
XX (GENP-) GENPHARM INT INC.
XX Kay RM, Lonberg N;
XX WPI: 1997-235888/21.
XX Novel anti-CD4 antibody produced by transgenic mice - used in the
XX treatment of auto-immune disease etc.
XX Example 42; Page 260-262; 396pp; English.
XX A novel composition has been developed which comprises an immunoglobulin
XX (Ig) having an affinity constant (Ka) of at least 2 multiply
XX 1000000000 M-1 for binding to a predetermined human antigen. The
XX present sequence represents the kappa light chain plasmid pCK7-96
XX which includes the kappa constant region and polyadenylation site. Anti-
XX CD4 antibodies may be used in therapeutic and diagnostic applications,
XX especially for the treatment of human diseases. These antibodies reduce
XX activity of CD4 cells and reduce undesirable autoimmune reactions,
XX inflammatory response and transplant rejection. Transgenic animals are
XX capable of producing heterologous antibodies of multiple isotypes by
XX undergoing isotype switching. These animals produce a first Ig type
XX that is necessary for antigen-stimulated B-cell maturation and can
XX switch to encode and produce one or more subsequent heterologous
XX isotypes.
XX Sequence 3881 BP; 981 A; 1014 C; 921 G; 965 T; 0 other;
XX
XX Query Match 100.0%; Score 1001; DB 18; Length 3881;
XX Best Local Similarity 100.0%; Pred. No. 4e-263;
XX Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 cctgcgaacttatccgcctccatccagctctatattgtgtgcgggaagctagatgaat 60
Db 1156 cctgcgaacttatccgcctccatccagctctatattgtgtgcgggaagctagatgaat 1215
QY 61 agtctgcagatlaatatgtttgcgaacgttgttgcattgtctaacaggaatcgttgtgtca 120
Db 1216 agtctgcagatlaatatgtttgcgaacgttgttgcattgtctaacaggaatcgttgtgtca 1275
QY 121 cgtctcgttgttgtatgtcttcatcagctcgcgtgtcccaagaatcagaagcgagttaaa 180
Db 1276 cgtctcgttgttgtatgtcttcatcagctcgcgtgtcccaagaatcagaagcgagttaaa 1335
QY 181 tgaatccccatgttgtgcaaaaagcggttagctctcgtctcgtctcgtctcgtctcgtcaga 240
Db 1336 tgaatccccatgttgtgcaaaaagcggttagctctcgtctcgtctcgtctcgtctcgtcaga 1395
QY 241 agtaagttgcccagctgttatcaactaatgttatacgttcgagcaactgataatcttact 300
Db 1396 agtaagttgcccagctgttatcaactaatgttatacgttcgagcaactgataatcttact 1455
QY 301 gtcatagcacgttaagatgtcttctgtgactggtgagtaactcaacacgaatctctga 360

```

```

Db 1456 gtcaagccatccgtaagaatgtcttctgtgactgtgtgacttaacaaagctatctga 1515
QY 361 gaatagtatgtcgccgacgaattgtctctgtcccgccggttaataaaggaataacccg 420
Db 1516 gaatagtatgtcgccgacgaattgtctctgtcccgccggttaataaaggaataacccg 1515
QY 421 ccacatagcagaactttaaaagtgctcatcatgtgaaacgttcttcggggcgaaatc 480
Db 1576 ccacatagcagaactttaaaagtgctcatcatgtgaaacgttcttcggggcgaaatc 1635
QY 481 tcaagatcttaacgcgtgttgagatccagttcgtatgaaccactctgtgacccaactga 540
Db 1636 tcaagatcttaacgcgtgttgagatccagttcgtatgaaccactctgtgacccaactga 1695
QY 541 tcttaagcatcttaacttaccacagcgttcttggtgagcaaaaacaggaaagcaaat 600
Db 1696 tcttaagcatcttaacttaccacagcgttcttggtgagcaaaaacaggaaagcaaat 1755
QY 601 gccgcaaaaaggaataaaggcgacacaggaatgttgaaactacactcttctt 660
Db 1756 gccgcaaaaaggaataaaggcgacacaggaatgttgaaactacactcttctt 1815
QY 661 caatatattgaacattatcaaggttatgtctcatgagcgagatcatatttgatgt 720
Db 1816 caatatattgaacattatcaaggttatgtctcatgagcgagatcatatttgatgt 1875
QY 721 attagaaaaataaacgaataagggttcgcgcacacattcccccgaagaagtcgac 780
Db 1876 attagaaaaataaacgaataagggttcgcgcacacattcccccgaagaagtcgac 1935
QY 781 gtccaaagaacacattatcatcatgaacttaacataaagaagcgatcacagagcc 840
Db 1936 gtccaaagaacacattatcatcatgaacttaacataaagaagcgatcacagagcc 1995
QY 841 ttctgtcgcgcgttccgtgtagatcgatgaacacactctgcaacatgagccgag 900
Db 1996 ttctgtcgcgcgttccgtgtagatcgatgaacacactctgcaacatgagccgag 2055
QY 901 acggtcacagcttctgttaagcggatgcggagcagacacagccctcagggcgctga 960
Db 2056 acggtcacagcttctgttaagcggatgcggagcagacacagccctcagggcgctga 2115
QY 961 gcgggtgttgcggtgtcggtggtggttaactatcgagc 1001
Db 2116 gcgggtgttgcggtgtcggtggtggttaactatcgagc 2156

```

## RESULT 15

AAZ21996  
ID AAZ21996 standard; DNA: 3881 BP.

XX AC AAZ21996;

XX DT 24-NOV-1999 (first entry)

XX DE Nucleotide sequence of plasmid pCK7-96.

XX KW Transgenic animal; heterologous antibody; hybridoma; B cell;  
KW transgenic mouse; human heavy chain transgene; digoxin;  
KW human light chain transgene; immortalized cell; immunoglobulin;  
KW Shiga-like toxin; autoimmune disease; cancer; infectious disease;  
KW transplant rejection; blood disorder; coagulation disorder; ss.

XX OS Synthetic.

XX PN MO9945962-A1.

XX PD 16-SEP-1999.

XX PF 12-MAR-1999; 99MO-US05535.

XX PR 13-MAR-1998; 98US-0042353.

```

XX (GENP-) GENPHARM INT INC.
XX
XX Lonberg N, Fishwild DM, Ball WJ;
XX WPI; 1999-551219/46.
XX
XX Novel transgenic non-human animals used to produce heterologous
XX antibodies
XX
XX Example 42: Page 311-313; 484pp; English.
XX
XX The specification describes transgenic animals that are capable of
XX producing a heterologous antibody. The antibodies are isolated from a
XX hybridoma, comprising B cells, that is obtained from a transgenic mouse
XX having a genome comprising a human heavy chain transgene and a human
XX light chain transgene. The B cells are used to immortalized cells
XX suitable for generating a hybridoma, which produces a detectable
XX amount of an immunoglobulin that specifically binds digoxin or
XX Shiga-like toxin. B cells from transgenic animals can be used to
XX generate hybridomas expressing monoclonal high affinity human sequence
XX antibodies. Antibodies produced from the transgenic animals of the
XX invention can be used to treat human diseases, e.g. autoimmune
XX diseases, cancer, infectious disease, transplant rejection, blood
XX disorders such as coagulation disorders and other diseases. The
XX present sequence is used in the course of the invention.
XX
XX Sequence 3881 BP; 981 A; 1014 C; 921 G; 965 T; 0 other;

```

Query Match 100.0%; Score 1001; DB 20; Length 3881;  
Best Local Similarity 100.0%; Pred. No. 4e-263;  
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

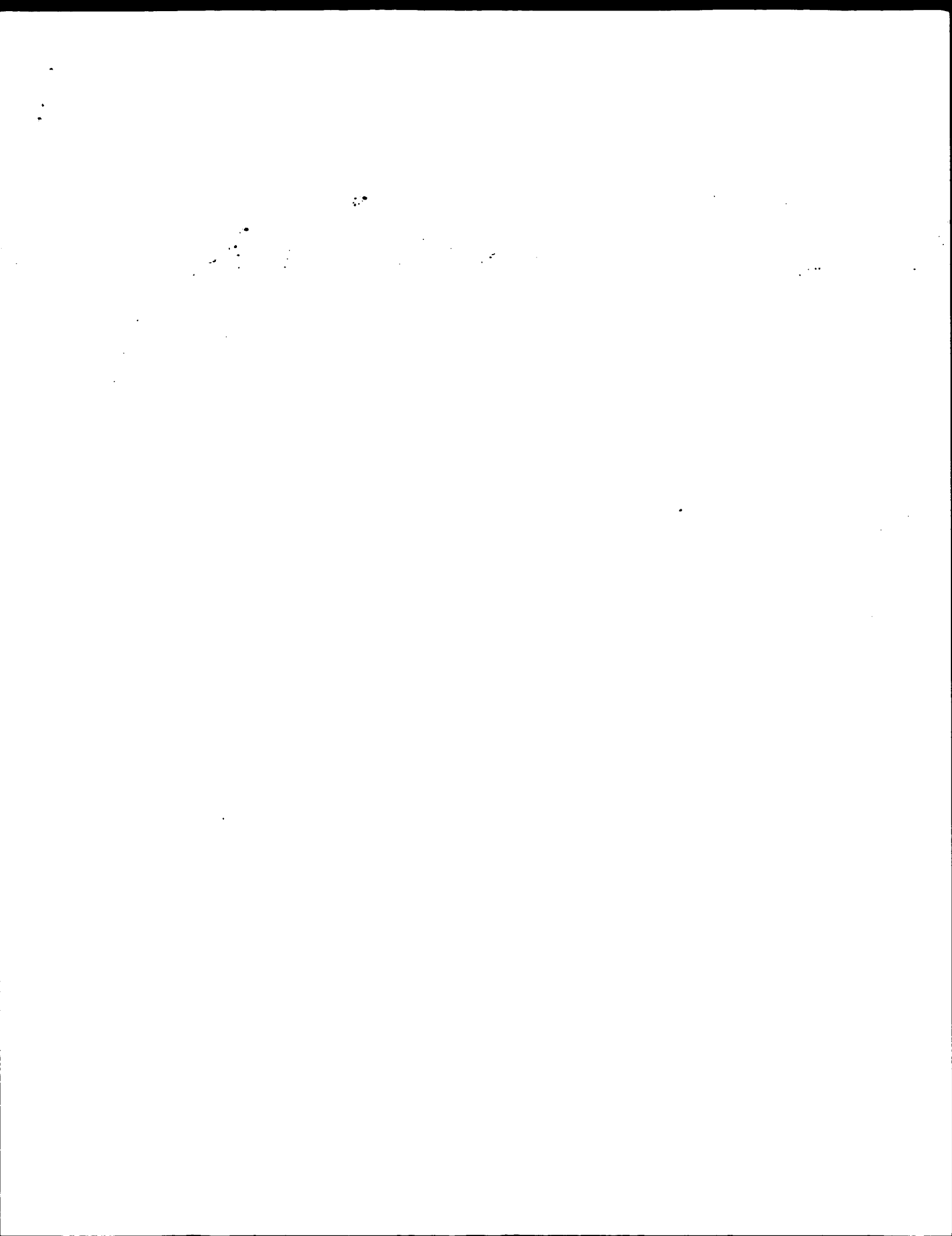
```

QY 1 cctgacattatccgcccacacagttcatatttcttcgggaagcctagatga 60
Db 1156 cctgacattatccgcccacacagttcatatttcttcgggaagcctagatga 1215
QY 61 agtcgccaatgaatgatttgcgaacgttctgtccatgtcacagcagtcgtgta 120
Db 1216 agtcgccaatgaatgatttgcgaacgttctgtccatgtcacagcagtcgtgta 1275
QY 121 cgtcgtgttggataggttcatcagctcgggttcccaagatcagagttca 180
Db 1276 cgtcgtgttggataggttcatcagctcgggttcccaagatcagagttca 1335
QY 181 tgatccccatgttgcgaaaaacggttagctccttcggtccctcagatgttga 240
Db 1336 tgatccccatgttgcgaaaaacggttagctccttcggtccctcagatgttga 1395
QY 241 agtaagttggccgacgtgttatcatcatgttatgagcagcatgacataatcttact 300
Db 1396 agtaagttggccgacgtgttatcatcatgttatgagcagcatgacataatcttact 1455
QY 301 gtcatgcatccgttaagatgcttctctgtgactgtgtgactcaacaatcatctga 360
Db 1456 gtcatgcatccgttaagatgcttctctgtgactgtgtgactcaacaatcatctga 1515
QY 361 gaatagttatgtcgccgacgaattgtctctgcccggcgtaataacggataacccg 420
Db 1516 gaatagttatgtcgccgacgaattgtctctgcccggcgtaataacggataacccg 1575
QY 421 ccacatagcagaactttaaaagtgctcatcatgtgaaacgttcttcggggcgaaatc 480
Db 1576 ccacatagcagaactttaaaagtgctcatcatgtgaaacgttcttcggggcgaaatc 1635
QY 481 tcaagatcttaacgcgtgttgagatccagttcgtatgaaccactctgtgacccaactga 540
Db 1636 tcaagatcttaacgcgtgttgagatccagttcgtatgaaccactctgtgacccaactga 1695
QY 541 tcttaagcatcttaacttaccacagcgttcttggtgagcaaaaacaggaaagcaaat 600
Db 1696 tcttaagcatcttaacttaccacagcgttcttggtgagcaaaaacaggaaagcaaat 1755

```

QY 601 gccgcaaaaaagggaataaggcgacacggaatgttgaatactactcttctt 660  
|||||  
Db 1756 gccgcaaaaaagggaataaggcgacacggaatgttgaatactactcttctt 1815  
|||||  
QY 661 caatatattgaagcaattatcaagggtatgtcctcagcgagacatatattgaatgt 720  
|||||  
Db 1816 caatatattgaagcaattatcaagggtatgtcctcagcgagacatatattgaatgt 1875  
|||||  
QY 721 attagaataataacaataagggttcgcgcacattcccgaaaagtgcacactgac 780  
|||||  
Db 1876 attagaataataacaataagggttcgcgcacattcccgaaaagtgcacactgac 1935  
|||||  
QY 781 gtctaagaaacattatcatgacatttaacctataaaataggcgta tcaagagccc 840  
|||||  
Db 1936 gtctaagaaacattatcatgacatttaacctataaaataggcgta tcaagagccc 1995  
|||||  
QY 841 ttctgtctcgcggttctcggtgatgaagtgaaaacctctgacacatgcagctcccgag 900  
|||||  
Db 1996 ttctgtctcgcggttctcggtgatgaagtgaaaacctctgacacatgcagctcccgag 2055  
|||||  
QY 901 acggtacaagctgtctgttaagcgga tgcgggagcagaacaagcccgtaaggcggtca 960  
|||||  
Db 2056 acggtacaagctgtctgttaagcgga tgcgggagcagaacaagcccgtaaggcggtca 2115  
|||||  
QY 961 gcgggtgttgcgcggtcgcggtcggttaactatgcgc 1001  
|||||  
Db 2116 gcgggtgttgcgcggtcgcggtcggttaactatgcgc 2156  
|||||

Search completed: January 8, 2002, 17:27:47  
Job time: 12647 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 8, 2002, 17:19:07 ; Search time 177.22 Seconds  
(without alignments)  
1279.226 Million cell updates/sec

Title: US-09-740-211-13\_COPY\_6000\_7000

Perfect score: 1 ccgtgcaacttaccgcctc.....gctgctgaactatgcgcgc 1001

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

#### Database :

Issued Patents NA:  
1: /cgn2\_6/prodata/2/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/prodata/2/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/prodata/2/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/prodata/2/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/prodata/2/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/prodata/2/ina/backfilest1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1001	100.0	3104	1	US-07-415-307A-1
2	1001	100.0	3104	1	US-08-371-320-1
3	1001	100.0	3343	6	5453363-2
4	1001	100.0	3819	4	US-09-042-353-393
5	1001	100.0	3819	4	US-08-758-417A-243
6	1001	100.0	3875	4	US-09-039-982A-21
7	1001	100.0	3875	4	US-09-039-641-21
8	1001	100.0	3875	4	US-09-039-762A-21
9	1001	100.0	3878	4	US-09-039-982A-27
10	1001	100.0	3878	4	US-09-039-641-27
11	1001	100.0	3878	4	US-09-039-762A-27
12	1001	100.0	3881	4	US-08-042-353-369
13	1001	100.0	3881	4	US-08-758-417A-217
14	1001	100.0	3883	4	US-09-039-982A-30
15	1001	100.0	3883	4	US-09-039-641-30
16	1001	100.0	3883	4	US-09-039-762A-30
17	1001	100.0	3908	4	US-09-039-982A-24
18	1001	100.0	3908	4	US-09-039-641-24
19	1001	100.0	3908	4	US-09-039-762A-24
20	1001	100.0	4045	4	US-08-464-700-54
21	1001	100.0	4273	3	US-08-795-430-3
22	1001	100.0	4622	4	US-08-509-024-6
23	1001	100.0	4622	4	US-09-333-279-6
24	1001	100.0	4696	2	US-08-929-967-15
25	1001	100.0	4723	4	US-09-042-353-370
26	1001	100.0	4723	4	US-08-758-417A-218
27	1001	100.0	4926	4	US-09-042-353-418

28	1001	100.0	4926	4	US-08-758-417A-268	Sequence 268, App
29	1001	100.0	4950	3	US-08-789-333F-58	Sequence 58, App1
30	1001	100.0	5158	2	US-08-929-967-16	Sequence 16, App1
31	1001	100.0	5181	3	US-08-801-344-5	Sequence 5, App1
32	1001	100.0	5181	4	US-09-498-399-5	Sequence 5, App1
33	1001	100.0	5292	2	US-08-793-610-3	Sequence 3, App1
34	1001	100.0	5294	2	US-08-793-610-2	Sequence 2, App1
35	1001	100.0	5323	2	US-08-793-610-1	Sequence 1, App1
36	1001	100.0	5364	2	US-08-793-610-4	Sequence 4, App1
37	1001	100.0	5639	4	US-09-175-690A-1	Sequence 1, App1
38	1001	100.0	5865	4	US-09-011-745-8	Sequence 8, App1
39	1001	100.0	6028	4	US-09-011-745-5	Sequence 5, App1
40	1001	100.0	6061	4	US-09-011-745-6	Sequence 6, App1
41	1001	100.0	6253	3	US-08-893-327-15	Sequence 15, App1
42	1001	100.0	6280	3	US-08-893-327-17	Sequence 17, App1
43	1001	100.0	6280	2	US-08-893-327-19	Sequence 19, App1
44	1001	100.0	6295	2	US-08-659-206A-4	Sequence 4, App1
45	1001	100.0	6306	5	PCT-US94-00658-1	Sequence 1, App1

#### ALIGNMENTS

RESULT 1  
US-07-415-307A-1

Sequence 1, Application US/07415307A  
Patent No. 5344757

#### GENERAL INFORMATION:

APPLICANT: Holte, Hans Joachim

APPLICANT: Seibl, Rudolf

APPLICANT: Schmitz, Gudrun

APPLICANT: Scholer, Hans

APPLICANT: Kessler, Christoph

APPLICANT: Maties, Ralf

TITLE OF INVENTION: Process for the Detection of Nucleic Acids

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Felte & Lynch

STREET: 805 Third Avenue

CITY: New York City

STATE: New York

COUNTRY: USA

ZIP: 10022

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage

COMPUTER: IBM PS/2

OPERATING SYSTEM: PC/DOS

SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/415, 307A

FILING DATE: 19900109

CLASSIFICATION: 435

Prior Application DATA:

APPLICATION NUMBER: PCT/EP89/0026

FILING DATE: 12-Jan-1989

APPLICATION NUMBER: DE 38 00 642.1

FILING DATE: 12-Jan-1988

APPLICATION NUMBER: DE 38 13 278.8

FILING DATE: 20-Apr-1988

ATTORNEY/AGENT INFORMATION:

NAME: Hanson, No. 5344757man D.

REGISTRATION NUMBER: 30, 946

REFERENCE/DOCKET NUMBER: BOER 798

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-688-9200

TELEFAX: 212-838-3884

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 3104 base pairs

TYPE: NUCLEIC ACID

STRANDEDNESS: single

TOPOLOGY: linear

US-07-415-307A-1

Query Match 100.0%; Score 1001; DB 1; Length 3104;  
 Best Local Similarity 100.0%; Pred. No. 2, 1e-301;  
 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cctgcaacttaccgctccatccatccatctaatgttgcgggaagctagaagta 60  
 |||||  
 Db 1549 CCTGCACCTTATCCGCTCCATCCACTATATATTGTCGGGGAAGCTAGAGT 1608  
 |||||  
 QY 61 agtcgcaagtaataatgttgcgaacgttgttgccttctcaagacatcggtgtca 120  
 |||||  
 Db 1609 AGTTGCGCAGTTAATAGTTGCGAAGCTGTGCTCAATGCTACAGGCATCGTGCTCA 1668  
 |||||  
 QY 121 cgtctgctgttggatgatgcttccatccgctccgctcccaagcatcaagcgagttaca 180  
 |||||  
 Db 1669 CGCTGCTGCTTGTGATGSGCTTCAATTCAGCTCCGCTTCCACAGCATCAAGCGAGTTACA 1728  
 |||||  
 QY 181 tgatcccccatgtgtgcataaaagcggttagctccttcggctcccgatcgtgtcaga 240  
 |||||  
 Db 1729 TGATCCCCCATGTGTGCAAAAAAGCGGTAGCTCCCTCCGCTCCGATCGTTGTCAGA 1788  
 |||||  
 QY 241 agtaagttggccgcaagtgtatcatcatatgtatggcaacactgtatcttact 300  
 |||||  
 Db 1789 AGTAAGTTGGCCGCGAGTGTATCACTATGTTATGGACACTGCATATATTTCTTACT 1848  
 |||||  
 QY 301 gtcatacctcgttaagaatcttctgtgacgtgtgagtaactcaacaaatcatctga 360  
 |||||  
 Db 1849 GTCATGCCATCCGTAAGATGCTTTCTGTGACTGTGAGTACCAACCAAGTATCTGA 1908  
 |||||  
 QY 361 gaataagtgatggcgacccgaatgtctcttgcggcgatcaatacgagataaccgag 420  
 |||||  
 Db 1909 GAATAGTGTATGGCGGACGAGTGTCTTGGCCCGGCAATACGGGATTAATACCGG 1968  
 |||||  
 QY 421 ccaatagcagaacttaaaagtgctcatcatgtgaaagagttcttcgagggcaaaac 480  
 |||||  
 Db 1969 CCACATGAGCAACTTTAAAGTGCATCATATGGAAGAACGTTCTTCGGGCGCAAAAC 2028  
 |||||  
 QY 481 tcaagaacttaccgctgtgtgagatccagtgatgaaacacatcgttgcacccaatga 540  
 |||||  
 Db 2029 TCAAGGATCTTACCGCTGTGAGATCCAGTTCATGTAACCCACTCGTGCACCAACTGA 2088  
 |||||  
 QY 541 tcttcagcaactttgttccaccagtttctgggtgtgagcaaaaaacaggaaggaat 600  
 |||||  
 Db 2089 TCTTTCAGCATCTTTACTTTCACACAGGTTGCGGTGAGCAAAACAGGAAGGCAAAAT 2148  
 |||||  
 QY 601 gccgcaaaaaggaaataaggcgacacgaaatgttgaataactcaactcttctt 660  
 |||||  
 Db 2149 GCCGCAAAAAAGGAAATAGGGGACACGAAATGTTGATACTCTTCTCTTTT 2208  
 |||||  
 QY 661 caatatattgaagcaatcaagggatattgtctcagaagcgagataacttgaatgt 720  
 |||||  
 Db 2209 CAATATATTGTAAGCATTTATCAGGGTTATGTCATAGCGGATACATATTTTGAATGT 2268  
 |||||  
 QY 721 attagaataataaacaataaggggttcgagcaacttcccgaaagatgcacactga 780  
 |||||  
 Db 2269 ATTATAGAAAAATTAACAAATAGGGGTTCCGGCACATTTCCCGAAGATGCACTGTGAC 2328  
 |||||  
 QY 781 gtctaaagaacattatcatgacatataactataaataagcgatcagagcc 840  
 |||||  
 Db 2329 GUCTAGAAACCATATATATCATGACATTAACCTAATAAATAGGGTATCAGAGGCC 2388  
 |||||  
 QY 841 tttcgtctcgcggttgcgtgatgagcggtgaagaaccttgacacatcgactcccgag 900  
 |||||  
 Db 2389 TTTTCGTCTGCGGCTTTCGGGTATGACGGTGAAGAACCTTGTGACATGACGCTCCCGAG 2448  
 |||||  
 QY 901 acggtacagctgtctgtaagcgagtcgaggaagcaagcccggtcagggcggtca 960  
 |||||  
 Db 2449 ACGGTACAGCTTGTCTGTAAGCGATTCGGGAGAGCAAGCCCGTCAAGGCGCGTCA 2508  
 |||||  
 QY 961 gcgagtgctgagcggtgtcggggtgtgacttaactatgcggc 1001  
 |||||  
 Db 2509 GCGGGTGTTGGCGGCTGTGCGGGCTGGCTTAACATGCGGC 2549

RESULT 2  
 US-08-371-320-1  
 ; Sequence 1, Application US/08371320  
 ; Patent No. 5702888  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Holthe, Hans Joachim  
 ; APPLICANT: Selbl, Rudolf  
 ; APPLICANT: Schmitz, Gudrun  
 ; APPLICANT: Scholler, Hans  
 ; APPLICANT: Kessler, Christoph  
 ; APPLICANT: Matles, Ralf  
 ; TITLE OF INVENTION: Process for the Detection of Nucleic Acids  
 ; NUMBER OF SEQUENCES: 3  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Felte & Lynch  
 ; STREET: 805 Third Avenue  
 ; CITY: New York City  
 ; STATE: New York  
 ; COUNTRY: USA  
 ; ZIP: 10022  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage  
 ; COMPUTER: IBM PS/2  
 ; OPERATING SYSTEM: PC/DOS  
 ; SOFTWARE: Wordperfect  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/371,320  
 ; FILING DATE: 11-JAN-1995  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 07/415,307  
 ; FILING DATE: 09-Jan-1990  
 ; APPLICATION NUMBER: PCT/EP99/0026  
 ; FILING DATE: 12-Jan-1999  
 ; APPLICATION NUMBER: DE 38 00 642.1  
 ; FILING DATE: 12-Jan-1998  
 ; APPLICATION NUMBER: DE 38 13 278.8  
 ; FILING DATE: 20-Apr-1998  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Hanson, No. 5702888man D.  
 ; REGISTRATION NUMBER: 30,946  
 ; REFERENCE/DOCKET NUMBER: BOER 798  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 212-688-9200  
 ; TELEFAX: 212-638-3884  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 3104 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: Linear  
 ; US-08-371-320-1

Query Match 100.0%; Score 1001; DB 1; Length 3104;  
 Best Local Similarity 100.0%; Pred. No. 2, 1e-301;  
 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cctgcaacttaccgctccatccatccatctaatgttgcgggaagctagaagta 60  
 |||||  
 Db 1549 CCTGCACCTTATCCGCTCCATCCACTATATATTGTCGGGGAAGCTAGAGT 1608  
 |||||  
 QY 61 agtcgcaagtaataatgttgcgaacgttgttgccttctcaagacatcggtgtca 120  
 |||||  
 Db 1609 AGTTGCGCAGTTAATAGTTGCGAAGCTGTGCTCAATGCTACAGGCATCGTGCTCA 1668  
 |||||  
 QY 121 cgtctgctgttggatgatgcttccatccgctccgctcccaagcatcaagcgagttaca 180  
 |||||  
 Db 1669 CGCTGCTGCTTGTGATGSGCTTCAATTCAGCTCCGCTTCCACAGCATCAAGCGAGTTACA 1728  
 |||||  
 QY 181 tgatcccccatgtgtgcataaaagcggttagctccttcggctcccgatcgtgtcaga 240

Db 1729 TGAATCCCGCCATGTTGCAAAAAAGCGGTTAGCTCCCTCGGTCCTCGATCGTTGCA 1788  
OY 241 agtaagttgacgagtgatcatcaatbgtatgagcaactgacataatcttact 300  
Db 1789 AGTAAGTGGCCGACAGTGTATGATCATGATGATGACACACTGATATATCTTACT 1848  
OY 301 gtaatgacatcgttaagatgcttcttctgactcgtgagtaactcaacgaatcttca 360  
Db 1849 GTCAATGCAATCCGTAAGATGCTTTTCTGATGATGATGATGATGATGATGATGAT 1908  
OY 361 gaataagatgagcgagcgagcgagctgcttctgacgagcgagcaataagagtaatacagcg 420  
Db 1909 GAATAGTATGCGGACGACGAGTGTCTTGGCCGCGATACGAGTAATACCGCG 1968  
OY 421 ccacatgacgaactttaaagtgatcatcatgtaaaagcttcttctgagcgagaaactc 480  
Db 1969 CCACATAGCAAGACTTTTAAAGTGTCTCATGTAAGAAAGCTTCTTGGGCGCAAACTC 2028  
OY 481 tcaagagatctacgctgttgaagatccagttcagatgtaaccactcgtgacccaactga 540  
Db 2029 TCAAGAGATCTTACCGCTGTGAGATCCAGTTCGATGTAACCACTCGTGACCCCACTGA 2088  
OY 541 tcttcagatcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 600  
Db 2089 TCTTCAGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2148  
OY 601 gccgcaaaaaggaataaagggcgacacggaatgltgaaatactactacttcttctt 660  
Db 2149 GCCGCAAAAAAGGAGATAGGGCGACACGGAATGTTGAATACATCATCTTCTTCTT 2208  
OY 661 caataatattgaagcattatcaaggttatgtcttcttcttcttcttcttcttcttctt 720  
Db 2209 CAATATTTATTAACCATTTATCAAGGTTATGTCATGAGCGATACATATTTGAATGT 2268  
OY 721 attagaanaataaacaataaaggggttccgacacatttcccgcaaaaagtgacacactgac 780  
Db 2269 ATTATGAAAAATTAACAATAAGGGGTTCCGCGACATTTCCCGCAAAAGTGCACCTGAC 2328  
OY 781 gcttaagaacacattatatacatgacatttaactataaataaggtatcaagaagccc 840  
Db 2329 GTCATGAAGAACCATTTATATCATGACATTAACCTATTAATAAGGCGATACAGAGGCC 2388  
OY 841 tttcgttcgagcggttctgagatgagagtgaaacactcagacacagcagctcccgag 900  
Db 2389 TTTGCTCTCGGCGGTTTGGTATGATGACGCTGTAACCTCTGACACATGACACTCCCGAG 2448  
OY 901 acggtcacagcttctgtaagcgatgacgagagcgagacagacagccgtcaagcgagctca 960  
Db 2449 ACGGTACACAGCTTCTTGAAGCGATGCGGAGACAGACAGCCCGTACAGGCGCGTCA 2508  
OY 961 gcgaggttctgagcgaggttctgaggttctgaggttctgaggttctgaggttctgag 1001  
Db 2509 GCGGCTGTGGGCGGTGCTGCGGCGCTGCTTAACATATGCGGC 2549

## RESULT 3

5453363-2/c

Patent No. 5453363

APPLICANT: RUDOLPH, RAINER; FISCHER, STEPHAN; MATTES, RALF

TITLE OF INVENTION: PROCESS FOR THE ACTIVATION OF T-PA OR

ING AFTER GENETIC EXPRESSION IN PROKARYOTES

NUMBER OF SEQUENCES: 4

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/206,044

FILING DATE: 02-MAR-1994

PRIORITY APPLICATION DATA:

APPLICATION NUMBER: 942,370

FILING DATE: 09-SEP-1992

APPLICATION NUMBER: 498,500

FILING DATE: 23-MAR-1990

APPLICATION NUMBER: 76,207

FILING DATE: 23-OCT-1986

SEQ ID NO: 2:  
5453363-2  
LENGTH: 3343

Query Match 100.0%; Score 1001; DB 6; Length 3343;  
Best Local Similarity 100.0%; Pred. No. 2,2e-301;  
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 cctgcaacttaccgctccatccagctatatttcttgcgggaagcctagagtaagt 60  
Db 2188 CCTGCAACTTTATCCGCCATCCATCCAGTATTAATTTGTCGGGAAAGCTAGAGTAAGT 2129  
OY 61 agttccgcaatgaatgatttgcgaacgttctgacatgtcaagagcagtcgtgtgca 120  
Db 2128 AGTTCCCGATTAAATGTTGGCGCAACGTTGTCATTTGTCACAGGATGCTGTGCA 2069  
OY 121 cgtcgtcttctgataagcttcatcagctcagctcagcttcccaagatcaagcgagta 180  
Db 2068 CGTGTGCTTGGTATGCTTCAATTCACCTCCGCTTCCCAACGATCAAGCGGAGTTCACA 2009  
OY 181 tgatccccatgttctgcaaaaagcggttaagctcttcttcttcttcttcttcttcttct 240  
Db 2008 TGAATCCCGCATGTTGCAAAAAAGCGGTTACCTCTTCTGCTCCGATCGTTCTCA 1949  
OY 241 agtaagttgacgagtgatcatcaactaagtgatgacagcactgacataatcttact 300  
Db 1948 AGTAAGTGGCCGACAGTGTATACATCATGATGATGATGATGATGATGATGATGAT 1889  
OY 301 gtaatgacatccgtaagatgcttcttctgactgagtgactcaacgaatcattctga 360  
Db 1888 GTCATGCCATCCGTAGATAGCTTCTTCTGACTGTGAGTACTCAACCAAGTATCTGTA 1829  
OY 361 gaataagatgagcgagcgaggttcttcttccggcgctcaataagagataacggcg 420  
Db 1828 GAATAGTATGAGCGGACCGAGTCTCTTGGCGCGCTCAATAGGATATATACCGCG 1769  
OY 421 ccacatagcaaacctttaaagtgatcatcaatgtaaaagcttcttgcggcgcaaaactc 480  
Db 1768 CCACATAGCAAGACTTTAAAGTGTCAATCATGTAAGGAAAGCTTCTTCCGCGGAAACATC 1709  
OY 481 tcaagagatctacgctgttgaagatccagttcagatgtaaccactcgtgacccaactga 540  
Db 1708 TCAAGAGATCTTACCGGTGAGATCCAGTCCGATTAACCTCTGACACCCACTGA 1649  
OY 541 tcttcagatcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 600  
Db 1648 TCTTCAGCATCTTTTACTTTCACAGGTTCTTGGGTGAGCAAAACAGGAAGCAAAAT 1589  
OY 601 gccgcaaaaaggaataaagggcgacacaggaatgttgaatactactacttcttctt 660  
Db 1588 GCCGCAAAAAAGGAGATTAAGCGGACACGGAATGTTGAATACTCATCTTCTTCTT 1529  
OY 661 caataatattgaagcattatcaaggttattctgctcaatgagagatataatattgaatg 720  
Db 1528 CAATATTTATGAAGCATTTTACAGGTTATTTCTCATGACGCGATACATATTTGAATGT 1469  
OY 721 attagaanaataaacaataaaggttccgacacatttcccgaaagtccactgac 780  
Db 1468 ATTATGAAAAATTAACAATAAGGGGTTCCGCGACATTTCCGAAAGTGCACCTGAC 1409  
OY 781 gcttaagaacacattatatacatgacatttaactataaataaggtatcaagaagccc 840  
Db 1408 GTCATGAAGAACCATTTATATCATGACATTAACCTATTAATAAGGCGTATACAGAGGCC 1349  
OY 841 tttcgttcgagcggttctgagatgagagtgaaacactcagacacatgacagctcccgag 900  
Db 1348 TTTGCTCTCGGCGGTTTGGTATGATGAGGTAAGAACTCTGACACATGACGCTCCGAG 1289  
OY 901 acggtcacagcttctgtaagcgatgacgagagcgagacagacagccgtcaagcgagctca 960  
Db 1288 ACGGTACACAGCTTCTGTAAGCGGATCCGCGGAGCAGACAGCCCGTCAAGGCGGCTCA 1229

QY 961 ggcgggtgtgagcgaggtgctgagggcttgaactatgacgagc 1001  
Db 1228 GCGGGCTTTGGCGGGGTGTCGGGGCTTGAATATGATGCGGC 1188

## RESULT 4

US-09-042-353-393  
Sequence 393, Application US/09042353  
Patent No. 6255458  
GENERAL INFORMATION:  
APPLICANT: Lonberg, Nils  
TITLE OF INVENTION: Transgenic No. 6255458-Human Animals for  
NUMBER OF INVENTION: Producing Heterologous Antibodies  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/042,353  
FILING DATE: 13-MAR-1998  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/810,279  
FILING DATE: 17-DEC-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/853,408  
FILING DATE: 18-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/904,068  
FILING DATE: 23-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/990,860  
FILING DATE: 16-DEC-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/053,131  
FILING DATE: 26-APR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/096,762  
FILING DATE: 22-JUL-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/155,301  
FILING DATE: 18-NOV-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/161,739  
FILING DATE: 03-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/165,699  
FILING DATE: 10-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/209,741  
FILING DATE: 09-MAR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/352,322  
FILING DATE: 07-DEC-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/544,404  
FILING DATE: 10-OCT-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/728,463  
FILING DATE: 10-OCT-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US96/16433  
FILING DATE: 10-OCT-1996

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/758,417  
FILING DATE: 02-DEC-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US97/21803  
FILING DATE: 01-DEC-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph T.  
REGISTRATION NUMBER: 36,429  
REFERENCE/DOCKET NUMBER: 014643-009040US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 393:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3819 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
US-09-042-353-393

Query Match 100.0%; Score 1001; DB 4; Length 3819;  
Best Local Similarity 100.0%; Pred. No. 2.3e-301;  
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 1 cctgcaacttaccgcccaccacagctatattgttcgaggaagcagaagta 60  
Db 1156 CCTGCACCTTTATCCGCCCTCCAGTCVATTAATTTGTCGGGAAGCTAGAGT 1215  
QY 61 agtcgcaagtaataagtttgcgaacgtttgacattgttacaagcagtcgt 120  
Db 1216 AGTTCGCCAGTAATAATGTTGGCAAGCTGTGTCATTGCTACAGCATGTGTC 1275  
QY 121 cgtcgtcgttggatgagcttcaatcagctccggttcccaagcagtcag 180  
Db 1276 CGCTGCTGTTGATGATGCTTCATTCAGCTCCGGTCCCAACATCAAGCGAGT 1335  
QY 181 tgaatcccaatgttgcgaanaaagcgttgccttcggttcctcagtcgtt 240  
Db 1336 TGATCCCCATGTTGCAAAAAAGCGTTAGCTTCCTTCGATCGATGTTGTC 1395  
QY 241 agtaagttgacgagtgatcaactcagtgatgagcagcagtcgaatctt 300  
Db 1396 AGTAAGTTGGCCGCGAGTGTATCACTCATGCTATGAGCAGCATGATATCTT 1455  
QY 301 gtaagcagtcgtaagaagtgctttctgagtgagtgagtaactcaacga 360  
Db 1456 GTCATGCCATCCGTAAGATGCTTTCTGTGACTGTGATCAACCAAGTCATTT 1515  
QY 361 gaatagtgatgcgagcagagtgcttcttgcgagcgatcaatacggagc 420  
Db 1516 GAATAGTGTATGCGCGAGCAGAGTTGCTTCCGCCGCGTCAATACGGATAT 1575  
QY 421 ccaatagcagaacttaaaagtgcatacaatggaacagcttctcggggcga 480  
Db 1576 CCACATACAGAACTTTAAAGTCTCATCTATGGAAGAGTTCTTGGGGCG 1635  
QY 481 tcaaggaatcctacgcgctgtgagatcagatcagatgaaacacactgtga 540  
Db 1636 TCAAGGATCTTACCGCTGTTGAGATCCAGTTGATGAAACCACTCTGACCA 1695  
QY 541 tctcagcatcttacttcaacagcgcttctggtgagcaaaaacaggaagca 600  
Db 1696 TCTTCAGATCTTTACTTTCACACAGGTTTCTGAGACAAAACAGCAAGCAAA 1755  
QY 601 ggcgcaaaaaggaataaaggcgacacggaataatgtaactactctctctt 660  
Db 1756 GCCGCAAAAAGGGAATAAGGCGCACGGAATAATGTTGATCTACTACTCTT 1815  
QY 661 caatatatgaagcattatcagggatattgtctatagcgagatataattg 720

Db 1816 CAATATTATGAGCATTTATCAGGCTTATGTCATGAGCGGATACATATTGAGT 1875  
 QY 721 atttagaataaataaagaaggttcgcgcacattcccgaaagtgccacgtac 780  
 Db 1876 ATTGAAAAATTAACAAATAGGGGTTCCGCGACATTTCCCGAAATGCGACCTGAC 1935  
 QY 781 gtctagaacacattatcagacattaacctataaataagcggtcagagccc 840  
 Db 1936 GTCTAGAAACCATTTATCATGACATTAACTATAAATAAGCGGTATCAGAGGCC 1995  
 QY 841 ttctgcgtcgcgttggtgltgagcgttgaaacctctgacacatgacgtccggag 900  
 Db 1996 TTTGCTCTCGCGCTTGGTGTGATGAGGATGAGAACCTCTGACACATGCACTCCGGAG 2055  
 QY 901 acggtcacactgtctgttaagcagatgacgcgagacagacccgctcaagcgctca 960  
 Db 2056 ACGGTCACTGTCTGTGTAAGGAGATGCGGAGACAGAACCCGTCAGGCGCGTCA 2115  
 QY 961 gcgggtgttgccggtgtcgaggctggttaactatgctgcgc 1001  
 Db 2116 GCGGCTGTGGCGGTGTCTCGGGCTGCTTAATGCTAGCGGC 2156

RESULT 5  
 US-08-758-417A-243  
 Sequence 243, Application US/08758417A  
 Patent No. 6300129

## GENERAL INFORMATION:

APPLICANT: Lomborg, Nils  
 Key, Robert M.

TITLE OF INVENTION: Transgenic No. 6300129-Human Animals for  
 Producing Heterologous Antibodies

NUMBER OF SEQUENCES: 417

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP  
 STREET: Two Embarcadero Center, Eighth Floor.  
 CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA: US/08/758,417A

FILING DATE: 02-Dec-1996

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/728,463

FILING DATE: 10-OCT-1996

APPLICATION NUMBER: US 08/544,404

FILING DATE: 10-OCT-1995

APPLICATION NUMBER: US 08/352,322

FILING DATE: 07-DEC-1994

APPLICATION NUMBER: US 08/209,741

FILING DATE: 09-MAR-1994

APPLICATION NUMBER: US 08/165,699

FILING DATE: 10-DEC-1993

APPLICATION NUMBER: US 08/161,739

FILING DATE: 03-DEC-1993

APPLICATION NUMBER: US 08/155,301

FILING DATE: 18-NOV-1993

APPLICATION NUMBER: US 08/096,762

FILING DATE: 22-JUL-1993

APPLICATION NUMBER: US 08/053,131

FILING DATE: 26-APR-1993

APPLICATION NUMBER: US 07/990,860

FILING DATE: 16-DEC-1992

ATTORNEY/AGENT INFORMATION:

NAME: Serafini, Andrew T.

REGISTRATION NUMBER: 41,303

REFERENCE/DOCKET NUMBER: 014643-00903005  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 576-0200  
 TELEFAX: (415) 576-0300  
 INFORMATION FOR SEQ ID NO: 243:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 3819 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA  
 SEQUENCE DESCRIPTION: SEQ ID NO: 243:  
 US-08-758-417A-243

Query Match 100.0%; Score 1001; DB 4; Length 3819;  
 Best Local Similarity 100.0%; Pred. No. 2,3e-301;  
 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cctgaacttatacgcgcctccacatcagatcattatgttgccgggaagctagaagt 60  
 Db 1156 CCTGCACTTATATCCGCTCCATCCAGTCTATTATGTTGCGGGAAGCTAGTAACT 1215  
 QY 61 agtgcgcagtaaatagtttgcgaagttgttgccattgtacagagcatcgtgtgtca 120  
 Db 1216 AGTTGCGCAGTTAATAGTTTGGCAACGTTTGGCATGTGACAGCATCGTGTGTCA 1275  
 QY 121 cgtcgtgtgttggatgagttcattcagtcgcgttcccaacagatcaagcgagttaca 180  
 Db 1276 CCTGCTGTTGGTATGAGCTTCAATTCAGTCCGGTCCCAAGATCAAGGCAATTTACA 1335  
 QY 181 tgatcccatgttgcgaacaaacggttagctcctcctcgtcctcagtcgtgtgtcaga 240  
 Db 1336 TGATCCCATGTGTGTGCAAAAAGGGGTAGCTCTTGTCGTCCTGTCAGTA 1395  
 QY 241 agtaagtgcgcagtggttatacctaagttatgagcagacatcattatccttact 300  
 Db 1396 AGTAACTGGCCGCACTGTATACATCATGATGATGAGCAGCACTGCATATTTCTTACT 1455  
 QY 301 gtcatgccatccgttaagatgtcttctgtgactgtgtgactgtgactgtgactgtgact 360  
 Db 1456 GTATATCCATCCGTAAAGATGCTTTTGTGACTGTGTGACTGTGACTGTGACTGTGACT 1515  
 QY 361 gaatagtatgagcgcagcagagttgtccttgcgcgggtcacaacagagataaccgct 420  
 Db 1516 GAATAGTATGAGCGGACGACGAGTGTCTTGGCCGCTCAATACGGGATATACCGCG 1575  
 QY 421 ccaatagcagaactttaaagtgtccatcatttggaacaaagtcttcggggggaacac 480  
 Db 1576 CCACATAGCAGCACTTTAAAGTCTATCATTTGAAAGCTTCTTCCGGGCGAAACTC 1635  
 QY 481 tcaagatcttaccgctgtgtgagatccagttcagatgtaaacacacgtgtcacccaactga 540  
 Db 1636 TCAGAGTCTTACCGCTGTGAGATCAGTTCAGTGTAAACCACTGTCCACCACTGA 1695  
 QY 541 tcttcagcalttacttaccacagcgttcttgggtgagcaaaacaggaaggaacaaat 600  
 Db 1696 TCTTACACATCTTTACTTTCACACAGCTTCTTGGGTGAGCAAAAACAGGAAGCAAAAT 1755  
 QY 601 gcgcgcaaaaggaataagggcgacacaggaagaaagtgtgaatactactcttctt 660  
 Db 1756 GCGCGCAAAAAGGAATAAGGCGACACGGAATGTGGAATACTCATCTTCTCTTTT 1815  
 QY 661 caatatatgaagcatllatcagggltatgttcacatgagcggatatacatlattaagt 720  
 Db 1816 CAATATTATGAGCATTTATCAGGCTTATGTCATGAGCGATACATATTGGAATG 1875  
 QY 721 atttagaataaataaagaaggttcgcgcacattcccgaaagtgccacgtac 780  
 Db 1876 ATTGAAAAATTAACAAATAGGGGTTCCGCGACATTTCCCGAAATGCGACCTGAC 1935  
 QY 781 gtctagaacacattatcagacattaacctataaataagcggtcagagccc 840

Db 1936 GTCTAAGAACATTATATACATTAACCTATAAAGGCTATACAGAGCCC 1995  
QY 841 ttctgtctgcgcgtttcgtgtatgacgltgaaacacctctacacatgcagctccgcgag 900  
Db 1996 TTTGGTCTCGCGGCTTTCGGTGGATGACGGTGAAGAACTCTGACACATGCAAGCTCCGGAG 2055  
QY 901 acggtcacagcttctgttaagcgatgcgagagcagacaagccgctcagcgcgctca 960  
Db 2056 ACGGTCAACAGCTTGTCTGTAAAGCGGATCCGGAGCAGACAAGCCGCTCAGGCGCGTCA 2115  
QY 961 gcgggtgtgcgcggtgtgcgaggtgcttaactatcagcgc 1001  
Db 2116 GCGGGTGTGGCGGCTGTCGGGCGCTGCTTAACATATCGCGC 2156

RESULT 6  
US-09-039-982A-21/c  
Sequence 21, Application US/09039982A  
Patent No. 6225042  
GENERAL INFORMATION:  
APPLICANT: Cal, Zeling  
APPLICANT: Sprent, Jonathan  
APPLICANT: Brunmark, Anders  
APPLICANT: Jackson, Michael  
APPLICANT: Peterson, Per A  
TITLE OF INVENTION: ANTIGEN PRESENTING SYSTEM AND METHODS FOR ACTIVATION OF T-CELL  
NUMBER OF SEQUENCES: 59  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Olson & Hierl, Ltd.  
STREET: 20 No. 6225042th Wacker Drive, Suite 3000  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/039,982A  
FILING DATE: 16-MAR-1998  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Olson, Arne M.  
REGISTRATION NUMBER: 20,203  
REFERENCE/DOCKET NUMBER: TSRI4710  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 580-1180  
TELEFAX: (312) 580-1189  
INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3875 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHEICAL: NO  
ANTI-SENSE: NO  
US-09-039-982A-21

Query Match 100.0% Score 1001; DB 4; Length 3875;  
Best Local Similarity 100.0%; Pred. No. 2,4e-301;  
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 cctgcacatttaccgctccatccagtcattatgttgcgaggaagcagagtaagt 60  
Db 2491 CCTGCAACTTATCCGCCATCCATGCTATATATGTTGCCGGAAGCTAAGT 2432  
QY 61 aattgcgcgttaataagtttgcgcaagctgttgcattgtacagcagatcgtgtgtca 120  
Db 2431 AATTGCGCAGTTAATAATGTTGCGCAACGTTGTTCCATTGCTACAGCATGCTGTGTC 2372

QY 121 cctgcgtctgttgatggcttcattcaagctccggttcccaacgatcaagcgagttaca 180  
Db 2371 CGCTGCTGTTTGATGAGCTTTCATTCACCTCCGGTCCCAAGATCAAGGAGTTTACA 2312  
QY 181 tgaatcccatgttctgcaaaaaaagcgltgaagctccttcgctccatcgtgtgtca 240  
Db 2311 TGATCCCATGTTTGCAAAAAAGGGTTACTCTCTTGCTCCGATCGCTTGTCA 2252  
QY 241 agtaagttgcccgcagtgcttcaactcaatgltatgcaagcaatgataatcttact 300  
Db 2251 AGTAAGTTTGCCGCACTGTTTACCTCATGCTTATGCGAGCACTGCATAATTTCTACT 2192  
QY 301 gtcattccatccgtaagatgcttctctgtgactgtgtgacttaacaaatcttctga 360  
Db 2191 GTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGTGACTCAACCAAGTATTCGA 2132  
QY 361 gaatagtatgcgcgcgagcaggtgtgcttctgcccgcgtcaatacagataacgcg 420  
Db 2131 GAATAGTGTATGCGCGAGCCAGTGTGCTTCCCGGCTCAATACGGGATTAATCCGCG 2072  
QY 421 ccacatagagaactttaaagtgctcaatcattgaaagcgttctcgggcgaaatc 480  
Db 2071 CCACATAGCAAGACTTTAAAGTCTCATATGGAAGAACGTTTCGGGCGAAGACTC 2012  
QY 481 tcaagatcttccgctgttgatccagttcgatgtaaccacatcgtaaccacatga 540  
Db 2011 TCAAGATCTTACCGCTGTGATGATCCAGTGTGATGATACCCACTCGTGCACCAACTGA 1952  
QY 541 tcttaagcatcttacttcaaccagcgttctgtgtgagcaaaaaacagaaagcaaat 600  
Db 1951 TCTTCAGCATCTTTTACTTTCACCGAGCGTTCTGGGTGAGCAAAAAACGAAAGCAAAAT 1892  
QY 601 gccgcaaaaaaggaataaggcgacaggaagtgtaataaccatacttcttctt 660  
Db 1891 GCCGCAAAAAAGGATTAAGGCGACAGGAATGTTGAATATCTATATCTTCTTTT 1832  
QY 661 caatattatgaagcattatcaaggttatgtctcatalgagcagatacatattgaaat 720  
Db 1831 CAATATATTTGAAGCATTTATCAGGTTATGCTCATGACAGGATACATATTTGATGT 1772  
QY 721 attagaataataaataaaggttccgcgcacattcccgaaagtgcaactgtac 780  
Db 1771 ATTAGAATAAATAAACAATAGGGGTTCCGGCACATTTCCCAAAAAGTCCACCTGTAC 1712  
QY 781 gtctaaagaacatlatatcatgacattaacctataaataaagtgatcaagagcc 840  
Db 1711 GTCTAAGAACCATTTATATCATGACATTAACCTATAAANTAGCGCTATCAGAGGCC 1652  
QY 841 ttctgtcgcgcgttctggtgatgcggtgaaacacttgcacatgcagctccgag 900  
Db 1651 TTTGCTTCGCGCGCTTTCGATGACGGTGAAACCTTGACACATGCAAGCTCCGAG 1592  
QY 901 acggtcacagcttctgttaagcgatgcgagagcagacaagccgctcagcgcgctca 960  
Db 1591 ACGGTCAACGCTTGTCTGTAAAGCGGATCCGGAGCAGACAAGCCGCTCAGGCGCTCA 1532  
QY 961 gcgggtgtgcgcggtgtgcgaggtgcttaactatcagcgc 1001  
Db 1531 GCGGGTGTGGCGGCTGTCGGGCGCTGCTTAACATATCGCGC 1491

RESULT 7  
US-09-039-641-21/c  
Sequence 21, Application US/09039641  
Patent No. 6251627  
GENERAL INFORMATION:  
APPLICANT: Cal, Zeling  
APPLICANT: Sprent, Jonathan  
APPLICANT: Brunmark, Anders  
APPLICANT: Jackson, Michael  
APPLICANT: Peterson, Per A  
TITLE OF INVENTION: ANTIGEN PRESENTING SYSTEM AND METHODS FOR

TITLE OF INVENTION: ACTIVATION OF T-CELLS  
 NUMBER OF SEQUENCES: 45  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Olson & Hierl, Ltd.  
 STREET: 20 No. 6251627th Wacker Drive, Suite 3000  
 CITY: Chicago  
 STATE: Illinois  
 COUNTRY: USA  
 ZIP: 60606  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/039,641  
 FILING DATE: 8-MAR-1995  
 CLASSIFICATION:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Olson, Arne M.  
 REGISTRATION NUMBER: 30,203  
 REFERENCE/DOCKET NUMBER: TSRI4710  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (312) 580-1189  
 TELEFAX: (312) 580-1189  
 INFORMATION FOR SEQ ID NO: 21:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 3875 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 US-09-039-641-21

Query Match 100.0%; Score 1001; DB 4; Length 3875;  
 Best local similarity 100.0%; Pred. No. 2,4e-301;  
 Matches 1001; Conservat 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cctgcaacttaccgctgttgagatccagttcgatgtaaccacactgaccccaactgaa 60  
 DB 2491 cctgcaacttaccgctgttgagatccagttcgatgtaaccacactgaccccaactgaa 2432  
 QY 61 agtgcgcaacttaccgctgttgagatccagttcgatgtaaccacactgaccccaactgaa 120  
 DB 2431 agtgcgcaacttaccgctgttgagatccagttcgatgtaaccacactgaccccaactgaa 2372  
 QY 121 cgtcgtcgctgttgagatccagttcgatgtaaccacactgaccccaactgaa 180  
 DB 2371 cgtcgtcgctgttgagatccagttcgatgtaaccacactgaccccaactgaa 2312  
 QY 181 tgatcccatgtgtgacagaaagcggttagctcctcgctcgcgtcgcgtgacga 240  
 DB 2311 tgatcccatgtgtgacagaaagcggttagctcctcgctcgcgtgacga 2252  
 QY 241 agtaagttgacgagatgttatacactaagttatgagcagcaactgataattcttact 300  
 DB 2251 agtaagttgacgagatgttatacactaagttatgagcagcaactgataattcttact 2192  
 QY 301 gtcacgcatccgtaagatccttctgtgacggtgagtagtaccacaaagcaactcctga 360  
 DB 2191 gtcacgcatccgtaagatccttctgtgacggtgagtagtaccacaaagcaactcctga 2132  
 QY 361 gaatagtgatgagcgagacgagttgctcttgcggcgatgacgaaggaataaccggcg 420  
 DB 2131 gaatagtgatgagcgagacgagttgctcttgcggcgatgacgaaggaataaccggcg 2072  
 QY 421 ccaatagcagaactttaaaagtctcattatgtaaaagttcttggggggaataaactc 480  
 DB 2071 ccaatagcagaactttaaaagtctcattatgtaaaagttcttggggggaataaactc 2012

QY 481 tcaagatcttaccgctgttgagatccagttcgatgtaaccacactgaccccaactgaa 540  
 DB 2011 tcaagatcttaccgctgttgagatccagttcgatgtaaccacactgaccccaactgaa 1952  
 QY 541 tcttcaacatcttacttccacagcgttcttggtgagcaaaaacaggaaggaataat 600  
 DB 1951 tcttcaacatcttacttccacagcgttcttggtgagcaaaaacaggaaggaataat 1892  
 QY 601 gccgcaaaaaggaataagggcgacagcgaataatgttgaatactcattccttctt 660  
 DB 1891 gccgcaaaaaggaataagggcgacagcgaataatgttgaatactcattccttctt 1832  
 QY 661 caatattatgaacattatcagaaggtattgtctcagcagcagatattgaatgt 720  
 DB 1831 caatattatgaacattatcagaaggtattgtctcagcagcagatattgaatgt 1772  
 QY 721 attagaaaaataacaataaggggttcgcgcacatttcccccgaagaatgccaactgac 780  
 DB 1771 attagaaaaataacaataaggggttcgcgcacatttcccccgaagaatgccaactgac 1712  
 QY 781 gctaaagaaccattatatacagatgaacataaaataagggatcagcagagccc 840  
 DB 1711 gctaaagaaccattatatacagatgaacataaaataagggatcagcagagccc 1652  
 QY 841 ttgcgtcgcgcgttctcgatgacggtgaaacacctctgacacatgacgtcccgag 900  
 DB 1651 ttgcgtcgcgcgttctcgatgacggtgaaacacctctgacacatgacgtcccgag 1592  
 QY 901 acggtacagcttctgtaagcgaatgacggaagcagaacagccgctcagggcgatca 960  
 DB 1591 acggtacagcttctgtaagcgaatgacggaagcagaacagccgctcagggcgatca 1532  
 QY 961 gcgagtggtgacggtgtcgcggggtggttaactatgagcgcgc 1001  
 DB 1531 gcgagtggtgacggtgtcgcggggtggttaactatgagcgcgc 1491

RESULT 8  
 US-09-039-762A-21/C  
 Sequence 21, Application US/09039762A  
 Patent No. 6255073  
 GENERAL INFORMATION:  
 APPLICANT: Cal, Zeling  
 APPLICANT: Sprent, Jonathan  
 APPLICANT: Brummark, Anders  
 APPLICANT: Jackson, Michael  
 APPLICANT: Peterson, Per A.  
 TITLE OF INVENTION: ANTIGEN PRESENTING SYSTEM AND METHODS  
 NUMBER OF SEQUENCES: 59  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Olson & Hierl, Ltd.  
 STREET: 20 No. 6255073th Wacker Drive, 36th Floor  
 CITY: Chicago  
 STATE: Illinois  
 COUNTRY: USA  
 ZIP: 60606  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/039,762A  
 FILING DATE: 16-MAR-1998  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: OLSON, Arne M.  
 REGISTRATION NUMBER: 30,203  
 REFERENCE/DOCKET NUMBER: TSRI 471.0 DIV.2  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (312) 580-1189  
 TELEFAX: (312) 580-1189

; INFORMATION FOR SEQ ID NO: 21:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 3875 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: cDNA  
 ; HYPOTHEICAL: NO  
 ; ANTI-SENSE: NO  
 ; US-09-039-762A-21

Query Match 100.0%; Score 1001; DB 4; Length 3875;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-301;  
 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cctgcaacttaccgctccatccagtcagtcataatgttgcgggaagctagaagaat 60  
 DB 2491 CCGCAACTTATCCCGCCATCCAGTCTATTATTGTTCCGGGAGAGTAACT 2432  
 QY 61 agtgcagcttaataagtttgcgcaagttgttgcattgtctacaggaatcgtgtgca 120  
 DB 2431 AGTTCGCCAGTTAATAGTTGCGCAAGCTGTGTCATTGTACAGGCAATCGTGGTGCA 2372  
 QY 121 cgtcgtgtgttgatagcttcaatcagctcgcgttcccaagatcaagcgagttaca 180  
 DB 2371 CGCTCGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2312  
 QY 181 tgatcccccattgtgcacaaagcggttagctccttcgctcccccagctgtgtgca 240  
 DB 2311 TGATCCCATGTTGCAAAAAGGCTTACTCTTCTGCTCCATCTCTTCTGCA 2252  
 QY 241 agtaagttgcgcagctgttcaactatgttataggcagcaatcattctact 300  
 DB 2251 AGTAAGTTGCGCGCACTGTATCAGCATGTTATGGCAGCACTCATTAATCTCTACT 2192  
 QY 301 gtcattccatcgtgaagatgcttcttctgtactgttgaactgaactcaagcaatcttga 360  
 DB 2191 GTCATGCCATCCGTAAGATGCTTCTGTGACTGTGTGACTGTGACTGTGACTGTGACT 2132  
 QY 361 gaatagtgatgcgcgcagcagatgtcttcttgcgcgcgtcaatacgggataataccgcg 420  
 DB 2131 GAATAGTGTATCGCGGAGCGAGTGTCTTCTGCGGCTCAATCGGATATATCCCG 2072  
 QY 421 ccacatagcagactttaaaatgtctatcatgtgaaaacgttcttgcggcgcaaacctc 480  
 DB 2071 CCACATAGCAGACTTTAAAGTGTCTATCATGTGAAAACTTCTGCGGGAATACTC 2012  
 QY 481 tcaagatcttaccgctgttgcagtcagctcagctgaatcagcaatcgcgcagcagcagc 540  
 DB 2011 TCAAGATCTTACCGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1952  
 QY 541 tcttaagatcttcttactctcagcagcttcttgcgggtgagcaaaaaacgaggaagcaaat 600  
 DB 1951 TCTTACGATCTTCTTACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1892  
 QY 601 gccgcaaaaaggaataagggcgacagcaagaaatgttgaatacctatcttctctttt 660  
 DB 1891 GCCGCAAAAAGGAAATAGGGCGACAGCAAGAAATGTAATATCTTCTTCTTCTTCTTCTT 1832  
 QY 661 caatatcttgaagcatctatcagtggttatgttcatgagcgagatataattgaatgt 720  
 DB 1831 CAATATTTTGAAGCATTTATCAGGTTATGTCTCATGACGAGCATATATTGGAATGT 1772  
 QY 721 attagaataaataaataagtggttccgcgcaatcttcccgcaaaagtgccactgac 780  
 DB 1771 ATTAGAATAAATAAATAAGGGGTTCCGGCACATTTCCCGAAAAAGTCCACCTGAC 1712  
 QY 781 gtctaaagaacattatattatcttgaacatacctataaaataagcgtatatacgaagccc 840  
 DB 1711 GTCTAAGAACATTATATTATGATGACATTAACCTATAAATAAGGCGATACACAGAGCCC 1652  
 QY 841 ttctgctcgcgcttctgctgagtgagcaaacctctgaacatgacagctcccgag 900

DB 1651 TTTCGCTCGCCGCTTTCGCTGATACGCTGAATAACCTCTGACATCAGCTCCGGAG 1592  
 QY 901 acggtcacagcttctctgtaagcgaatgcgcggagcagacaaagccgctcagggcgctca 960  
 DB 1591 ACGGTCAAGCTTGTCTGTAAGCGGATCCGGGAGACAGCAAGCCCGCTCAGCGCGCTCA 1532  
 QY 961 gcgggtgttgagggtgtgcggggcgctgaactatgagc 1001  
 DB 1531 GCGGTTGTTGGCGGCTGCGGCTGCTTAATATGCTGCGC 1491

RESULT 9  
 US-09-039-982A-27/C  
 ; Sequence 27, Application US/09039982A  
 ; Patent No. 6225042

; GENERAL INFORMATION:  
 ; APPLICANT: Cal, Zelling  
 ; APPLICANT: Sprent, Jonathan  
 ; APPLICANT: Brunmark, Anders  
 ; APPLICANT: Jackson, Michael  
 ; APPLICANT: Peterson, Per A  
 ; TITLE OF INVENTION: ANTIGEN PRESENTING SYSTEM AND METHODS FOR ACTIVATION OF T-C  
 ; NUMBER OF SEQUENCES: 59  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Olson & Hierl, Ltd.  
 ; STREET: 20 No. 6225042th Wacker Drive, Suite 3000  
 ; CITY: Chicago  
 ; STATE: Illinois  
 ; COUNTRY: USA  
 ; ZIP: 60606

; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/039, 982A  
 ; FILING DATE: 16-MAR-1998  
 ; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Olson, Arne M.  
 ; REGISTRATION NUMBER: 30,203  
 ; REFERENCE/DOCKET NUMBER: TSRI4710  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (312) 580-1180  
 ; TELEFAX: (312) 580-1189

; INFORMATION FOR SEQ ID NO: 27:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 3878 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: cDNA  
 ; HYPOTHEICAL: NO  
 ; ANTI-SENSE: NO  
 ; US-09-039-982A-27

Query Match 100.0%; Score 1001; DB 4; Length 3878;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-301;  
 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cctgcaacttaccgctccatccagtcagtcataatgttgcgggaagctagaagaat 60  
 DB 2494 CCGCAACTTATCCCGCTCCATCCAGTCTATTATTGTTCCGGGAGAGTAACT 2435  
 QY 61 agtgcagcttaataagtttgcgcaagttgttgcattgtctacaggaatcgtgtgca 120  
 DB 2434 AGTTCGCCAGTTAATAGTTGCGCAAGCTGTGTCATTGTACAGGCAATCGTGGTGCA 2375  
 QY 121 cgtcgtgtgttgatagcttcaatcagctcgcgttcccaagatcaagcgagttaca 180



Db 2374 CCCTGCTGTTGGATGAGCTTCATTCAGCTCCGGTCCACACATGACGAGGAGTTACA 2315  
Qy 181 tgatcccccatgttctgcaaaaaacggttagctctcctcggctccatcggtgtcaga 240  
Db 2314 TGATCCCCCATGTTGGCAAAAAAGCGGTAGCTCTCTGGCTCTCCGATCGGTTGTGAGA 2255  
Qy 241 agtaagttgcccagatgtatcaactcagtgatgagcagactgacatctctact 300  
Db 2254 AGTAAGTTGGCCGACAGTGTATCACTCACTGATGAGCAGCTCATTAATTTCTTACT 2195  
Qy 301 gtcacgcatccgtaagatgtctctgactgtgtgactgactcaaccaactcctcga 360  
Db 2194 GTCATGCCATCCGTAAAGATGCTTTCTGTGACTGGTGAAGTACTCAACCAAGCATCTGGA 2135  
Qy 361 gaatagtgatgagcagcagcagatgtctctcctccgctcacaatacgggataaccgcy 420  
Db 2134 GAATAGTGTAGCGGCGACCGAGTTCCTCTTCCGGCGCTCAATACGGGATTAATACCGCG 2075  
Qy 421 ccacatagcagaactttaaaagtgcacatcatltgaaaaacgtctctcggggcgaaaaactc 480  
Db 2074 CCACATAGCAGAACTTTAAAGGCTCATCATTTGAAAAAGCTTCTTCCGGGCGAAACCTC 2015  
Qy 481 tcaaggtcttaccgctgttgatgaltcagttcagatgtaaccactcgtccaccaactga 540  
Db 2014 TCAAGGATCTTACCGCTGTGAGATCACTGATGATTAACCACTGTCACCAACTGA 1955  
Qy 541 tcttcacatcttactctcaccagcgttctcgggtgagcaaaaaacgaggaagcaaat 600  
Db 1954 TCTTCACATCTTTTACTTTCACACAGGTTCTGCGGTGAGCAAAAAACAGAGGCAAAAT 1895  
Qy 601 gcgcgaataaagggaataagggcgacacggaatgttgatatactactctctctt 660  
Db 1894 GCCCAAAAAAGGGAATAGGCGCACAGGAAATGTAATCTCACTCTTCTCTTTT 1835  
Qy 661 caatattatgaagcattatcaaggtatgtctcagatgagcagatcacatattgaagt 720  
Db 1834 CAATATTTTAAGCATTTTATCAGGGTTATTTGCTCATGAGCGGATTAATTTGAATGT 1775  
Qy 721 attgaaaaataaacaataaggggttcgcgcacatctcccgaaaaagggcaccctgac 780  
Db 1774 ATTTGAAAAATTAACAATAGGCGTTCCGCGCATTTCCGGAAGGCGCACTGAC 1715  
Qy 781 gtctagaagaacattatcatatcatatcaactataaataagcgatcacagagccc 840  
Db 1714 GTTAAAGAACCATTTATATCATGACATTAACATATAAATAGGCTATCAGAGGCCCC 1655  
Qy 841 ttctgctcgcgcggttcgtgcatgacggtgaaaaactctgacacatgacgtccggag 900  
Db 1654 TTTGCTTCGCGCGCTTTCGGTGGATGACGCTGAAAACTGTGACACATGACGCTCCGGAG 1595  
Qy 901 acggtcaacagcttctcgtctgacgagtcgagggagcagacagccgctcagggcgctca 960  
Db 1594 ACGGTCACAGCTTGTCTGTAAAGCGGATCCGGAGACACAGCCCGTCAAGCGGCGCTCA 1535  
Qy 961 gcgggtgttggcggtgtgagggcggtgacttaactatgagcgc 1001  
Db 1534 GCGGCTGTGGCGGGGTGTGCGGGGTGCTTAACATGCGGCG 1494

RESULT 10  
US-09-039-641-27/c  
Sequence 27, Application US/09039641  
Patent No. 6251627  
GENERAL INFORMATION:  
APPLICANT: Cal, Zeling  
APPLICANT: Sprent, Jonathan  
APPLICANT: Brunmark, Anders  
APPLICANT: Jackson, Michael  
APPLICANT: Peterson, Per A  
TITLE OF INVENTION: ANTIGEN PRESENTING SYSTEM AND METHODS FOR  
ACTIVATION OF T-CELLS  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:

ADDRESS: Olson & Herl, Ltd.  
STREET: 20 No. 6251627th Wacker Drive, Suite 3000  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/039, 641  
FILING DATE: 8-MAR-1995  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Olson, Arne M.  
REGISTRATION NUMBER: 30, 203  
REFERENCE/DOCKET NUMBER: TSI14710  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 580-1180  
TELEFAX: (312) 580-1189  
INFORMATION FOR SEQ. ID NO.: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3878 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHEICAL: NO  
ANTI-SENSE: NO  
US-09-039-641-27

Query Match 100.0%; Score 1001; DB 4; Length 3878;  
Best Local Similarity 100.0%; Pred. No. 2.4e-301;  
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 cctgcaacttatacgcctccatccacagtcactataatgttgcggggaagcctgaggaagt 60  
Db 2494 CCTGCAACTTATCCGCTCCATCCAGCTATTAATGTTGCGGGAAGCTGAGTAAGT 2435  
Qy 61 agttcgcagttatagtttggcgaaggttgtagcactgtctacagcagcagtggtgta 120  
Db 2434 AGTTGCGCAGTTAATAGTTTGCGACGTTGTCATTTGCTTACAGSCATCGTGCTCA 2375  
Qy 121 cgtctcgttggtagtgcttcaatcagtcgcggttcccaagatcaagcgagttaca 180  
Db 2374 CCCTGCTGTTGGATGAGCTTCATTCAGCTCCGTTCCCAAGCATCAAGCGAGTTACA 2315  
Qy 181 tgatcccccatgttctgcaaaaaacggttagctctcctcggctccatcggtgtcaga 240  
Db 2314 TGATCCCCCATGTTGGCAAAAAAGCGGTAGCTCTCTGGCTCTCCGATCGGTTGTGAGA 2255  
Qy 241 agtaagttgcccagatgtatcaactcagtgatgagcagactgacatctctact 300  
Db 2254 AGTAAGTTGGCCGACAGTGTATCACTCACTGATGAGCAGCTCATTAATTTCTTACT 2195  
Qy 301 gtcacgcatccgtaagatgtctctgactgtgtgactgactcaaccaactcctcga 360  
Db 2194 GTCATGCCATCCGTAAAGATGCTTTCTGTGACTGGTGAAGTACTCAACCAAGCATCTGGA 2135  
Qy 361 gaatagtgatgagcagcagcagatgtctctcctccgctcacaatacgggataaccgcy 420  
Db 2134 GAATAGTGTAGCGGCGACCGAGTTCCTCTTCCGGCGCTCAATACGGGATTAATACCGCG 2075  
Qy 421 ccacatagcagaactttaaaagtgcacatcatltgaaaaacgtctctcggggcgaaaaactc 480  
Db 2074 CCACATAGCAGAACTTTAAAGGCTCATCATTTGAAAAAGCTTCTTCCGGGCGAAACCTC 2015  
Qy 481 tcaaggtcttaccgctgttgatgaltcagttcagatgtaaccactcgtccaccaactga 540  
Db 2014 TCAAGGATCTTACCGCTGTGAGATCACTGATGATTAACCACTGTCACCAACTGA 1955

```

QY 541 tcttcagcatttcttacttcaaccagcttctcgtgtgagcaaaacaggaagcaaat 600
Db 1954 TCTTCAGCATCTTTACTCTTACACAGCGCTTCTGGGAGACAAAACAGAGCAAAAT 1895
QY 601 gccgcaaaaagggaataaaggcgacacaggaataatgtgaatactactcttctt 660
Db 1894 GCCGCAAAAAGGAATAAGGCGACACGCAAAATGTTGAATCTCATCTCTCTCTT 1835
QY 661 caataattgaagcaattatcagggttattgtctatgaagcggaataatttgaat 720
Db 1834 CAATATATTGAAGCATTTATTCAGGGTTATGTCTCATAGCGGATCATATTTAATGT 1775
QY 721 attagaagaataaacaataagggttccgcgcacattcccgaaagtgccactgac 780
Db 1774 ATTTAGAAAATAAACAAAATAGGGGTTCCGCGCACATTTCCCGAAAGTGCCACTGAC 1715
QY 781 gtctaaagaacattatctatcatgacattactaataaataaggcgtatcatcagaagccc 840
Db 1714 GTCTAAGAAACCATTTATTTATCATGACATTACCTTAAAAAATAGCGATACAGAGGCC 1655
QY 841 ttctcgtcgcgcgttccgtgtagacagtgaaacactctgacatcagctcccgag 900
Db 1654 TTTCGTCGCGCGGTTTCGCTGATGACGGTGAAACCTCTGACATGACAGCTCCCGAG 1595
QY 901 acggtacagcgttctgtgaagcgtatgcgcggagcagacaagccgtcaggcgctca 960
Db 1594 ACGGTACAGCTGTGTGTGAACGGATGCGGAGACAGCAAGCCCGTACAGGCGCTCA 1535
QY 961 gcgggtgttgcgggtgtcgggagcgttactactcgtcggc 1001
Db 1534 GCGGCTGTGCGGGGTGTGCGGGCTGCGCTTAATCACTATGCGGC 1494

```

## RESULT 11

```

US-09-039-762A-27/c
Sequence 27, Application US/09039762A
Patent No. 6255073
GENERAL INFORMATION:
APPLICANT: Cal, Zelling
APPLICANT: Sprent, Jonathan
APPLICANT: Brumark, Anders
APPLICANT: Jackson, Michael
APPLICANT: Peterson, Per A.
TITLE OF INVENTION: ANTIGEN PRESENTING SYSTEM AND METHODS
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Olson & Hiern, Ltd.
STREET: 20 No. 6255073ch Wacker Drive, 36th Floor
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/039,762A
FILING DATE: 16-MAR-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OLSON, Arne M.
REGISTRATION NUMBER: 30,203
REFERENCE/DOCKET NUMBER: TSRI 471.0 DIV.2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 580-1180
TELEFAX: (312) 580-1189
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 3878 base pairs

```

```

TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cdna
HYPOTHEetical: NO
ANTI-SENSE: NO
US-09-039-762A-27

Query Match
Best Local Similarity 100.0%; Score 1001; DB 4; Length 3878;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ccggaacttaccgctccatccagttcatatattgttgcgggaagctagaagta 60
Db 2494 CCGGAACCTTATCCGCCCTCCATCCAGTCTATTAATTTGTCGGGGAAGCTAGAGT 2435
QY 61 agtccgaattatagtttgcgaacgttgttgcattgtcagaagcatctgtgtca 120
Db 2434 AGTTCGCAAGTTAATAGTTTGGCAACGTTGTGCAATTTGCTACAGGATCTGTG 2375
QY 121 cgtcgtcgttggtagtgccttcatcagctccagttcccaagataagcgagttaca 180
Db 2374 CGCTCGCTGTTGGTATGAGCTTCATTCAGCTCCGGTCCCAAGCATCAAGCGAGT 2315
QY 181 tgaatcccatgtttgcaaaaagcgttagctctcctcgttcccgatcgtgtcaga 240
Db 2314 TGATCCCATATTTGTGCAAAAAGCGGTAGCTCTCTGCTCCGATCGTGTGAGA 2255
QY 241 agtaagtgtgcgcagtglttactactatgtttagcagcagctgcaataattctact 300
Db 2254 AGTAAGTTGGCGCAGCTGTATCACTCATGTGTTATGGCAGCATGATTTCTTACT 2195
QY 301 gtcacgcatccgttaagttgtcttctgtgacgtgtgagtaactcaacaaatcactga 360
Db 2194 GTCATGCCATCCGTAAAGATGCTTCTGTGACTGTGATGACTCAACCAAGTATTTCTGA 2135
QY 361 gaatagtgtagcgcgcagcagagtgctcctcgtccgcgcgtcaataagggataacggc 420
Db 2134 GAATAGTGTATGCGCGCAGCAGAGTTGCTTCCGCGCGCATATACGGGATTTACCGCG 2075
QY 421 ccagatagcagaactttaaagggtcctaactatgtgaaacgttcttcggcggaacactc 480
Db 2074 CCATATACAGAACTTTAAAGTCTCATATGGAAGAGTCTTCCGCGCGCAAAATC 2015
QY 481 tcaagatcttaccgctgtgtgagatccagttcagttgagcaaaacaggaagcaaat 540
Db 2014 TCAAGATCTTACCGCTGTGAGATCCAGTTCCATGTGAACCCACTGTGACCCACTGA 1955
QY 541 tcttcagcatttcttacttcaaccagcttctcgtgtgagcaaaacaggaagcaaat 600
Db 1954 TCTTCAGCATCTTTACTCTTACACAGCGCTTCTGGGAGACAAAACAGAGCAAAAT 1895
QY 601 gccgcaaaaagggaataaaggcgacacaggaataatgtgaatactactcttctt 660
Db 1894 GCCGCAAAAAGGAATAAGGCGACACGCAAAATGTTGAATCTCATCTCTCTCTT 1835
QY 661 caataattgaagcaattatcagggttattgtctatgaagcggaataatttgaat 720
Db 1834 CAATATATTGAAGCATTTATTCAGGGTTATGTCTCATAGCGGATCATATTTAATGT 1775
QY 721 attagaagaataaacaataagggttccgcgcacattcccgaaagtgccactgac 780
Db 1774 ATTTAGAAAATAAACAAAATAGGGGTTCCGCGCACATTTCCCGAAAGTGCCACTGAC 1715
QY 781 gtctaaagaacattatctatcatgacattactaataaataaggcgtatcatcagaagccc 840
Db 1714 GTCTAAGAAACCATTTATTTATCATGACATTACCTTAAAAAATAGCGATACAGAGGCC 1655
QY 841 ttctcgtcgcgcgttccgtgtagacagtgaaacactctgacatcagctcccgag 900
Db 1654 TTTCGTCGCGCGGTTTCGCTGATGACGGTGAAACCTCTGACATGACAGCTCCCGAG 1595

```



Db 1756 GCGCAAAAAGGGAATAGGCGACACGAAATGTGAATACACTCTTCTTTT 1815  
Qy 661 caatatattgaagcattatcaggggtatgtctatgtgagggataacataatttgatgt 720  
Db 1816 CAATATTATTGAACATTATACAGGGTTATGTCTCATGTAGCGGAATACATATTGATGT 1875  
Qy 721 atttgaataaataacaataaggggttcgcgcgcacattcccccgaagaatgcacactgac 780  
Db 1876 ATTTAGAAAATAAACAAATAGGGGTTCCCGCACATTTCCCGAAATAGCCACCTGAC 1935  
Qy 781 gttctaaacacattatcatcagacattacataaataagcgatcacagagcc 840  
Db 1936 GTCTAGAAAACCAATTATATCATGACATTAACTATAAATAGCGGTATCAGAGGCC 1995  
Qy 841 ttctgcctcgcgcgttcgtgtagacggtgaaacaccttcgacacatgacgtcccgag 900  
Db 1996 TTTTCGCTCCCGGTTTCGGTGTATGACGGTGAACAACTCTGACACATGCAAGCTCCCGAG 2055  
Qy 901 acggtacagcgttctgttaagcgcgaltgcgcggagcagacaagcccgtaagcgcgta 960  
Db 2056 ACGGTACAGACTTGTCTTAAGCGGATGCCGGAGCAGACAGCCGTCAGGCGCGTCA 2115  
Qy 961 gcgggttcgagcggttcgcgcgcgttcgtaactatgagcg 1001  
Db 2116 GCGGCTTGTGCGGCTGTGCGGCTGCTTAATGATGCGGC 2156

## RESULT 13

US-08-758-417A-217

Sequence 217, Application US/08758417A

Patent No. 6300129

GENERAL INFORMATION:

APPLICANT: Lomborg, Nils

Key, Robert M.

TITLE OF INVENTION: Transgenic No. 6300129-Human Animals for

Producing Heterologous Antibodies

NUMBER OF SEQUENCES: 417

CORRESPONDENCE ADDRESSES:

ADDRESS: Townsend and Townsend and Crew LLP

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/758,417A

FILING DATE: 02-Dec-1996

CLASSIFICATION: &lt;Unknown&gt;

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/728,463

FILING DATE: 10-OCT-1996

APPLICATION NUMBER: US 08/544,404

FILING DATE: 10-OCT-1995

APPLICATION NUMBER: US 08/352,322

FILING DATE: 07-DEC-1994

APPLICATION NUMBER: US 08/209,741

FILING DATE: 09-MAR-1994

APPLICATION NUMBER: US 08/165,699

FILING DATE: 10-DEC-1993

APPLICATION NUMBER: US 08/161,739

FILING DATE: 03-DEC-1993

APPLICATION NUMBER: US 08/155,301

FILING DATE: 18-NOV-1993

APPLICATION NUMBER: US 08/096,762

FILING DATE: 22-JUL-1993

APPLICATION NUMBER: US 08/053,131

FILING DATE: 26-APR-1993

APPLICATION NUMBER: US 07/990,860

FILING DATE: 16-DEC-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Serafini, Andrew T.  
REGISTRATION NUMBER: 41,303  
REFERENCE/DOCKET NUMBER: 014643-009030US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 217:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3881 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 217:  
US-08-758-417A-217

Query Match 100.0%; Score 1001; DB 4; Length 3881;  
Best Local Similarity 100.0%; Pred. No. 2,4e-301;  
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 cctgaacttaccgcctccatccagctcatatgaattgtccggagagctagaagtaagt 60  
Db 1156 CCTGAACCTTATCCGCTCCATCCACTCTTATTTGTTGCGGGAAGCTAGAGTAAGT 1215  
Qy 61 agtcgcagcagtaagattgagcagcgttgtagcattgctacagagcagcgttgtagc 120  
Db 1216 AGTTGCGCAGTAATTAATTTGCGCAACGTTGTGCCATTGCTACAGGCATCGTGTGCA 1275  
Qy 121 cgcctgcgtttgttgaagctcatcagctccggtcccaagatcaagcgagttaca 180  
Db 1276 CGCTCGTCGTTGTGATAGGCTTCAATTCAGCTCCGATTCCCAAGATCAAGCGAGTTACA 1335  
Qy 181 tgatcccccagctgtgcaaaaaagcggttagctccttcgctccgcagatcgttgtagc 240  
Db 1336 TGATCCCCCAATGTTGTGCAAAAAAGCGGTTAGCTTCGTCGTCGTCGATCGTTGACGA 1395  
Qy 241 agtaagcttcgagcagcgttcatcagctcatcagctcatcagctcatcagctcatcag 300  
Db 1396 AGTAAGTTGCGCGCAGGTTATCATCTCATGTTATGACACACTGCATTAATTCCTTACT 1455  
Qy 301 gtcagccatccgtaagatcttcttcgtagcgtgtagtcaacccaagatcatcga 360  
Db 1456 GTCATGCCATCCGTAAGATCTTTCTGTGATCGTGTGATGAGTCAACCAAGTATTCTGA 1515  
Qy 361 gaatagctgtagcagcagcagcgttcttcgtagcgtgtagtcaacccaagatcatcga 420  
Db 1516 GAATAGTGTATGCGGCGACCGAGTCTCTTCCCGCGTCAATACGGGATTAATACCGCG 1575  
Qy 421 ccaatagcagaactttaaagtgctcatcatcagtaaaacgcttcgagcggaacac 480  
Db 1576 CCAATAGCGAGACTTTAAAGTGTCTCATCTATGGAAGAGTTCTTCCGGGCAAAACAC 1635  
Qy 481 tcaagatcttaccgctgttgagatccagctcagatgtaaacccaactcgtgcaccaactga 540  
Db 1636 TCAAGGATCTTACCGCTGTGTGAGATCCAGTTGATGTATACCACTGTCACCACTGA 1695  
Qy 541 tcttgcagatcttcttcacccagcgttcttggtgagcaaaaaagcggaagcaaat 600  
Db 1696 TCTTCAGCATCTTTAACTTCAACAGCGCTTCTGGGTGCAAAAAAGCAAGCAAAAT 1755  
Qy 601 gccgcaaaaaaggaataagggcgagcaagaaatgttgaatcatcactctctctctt 660  
Db 1756 GCGCAAAAAGGGAATAGGCGACACGGAATGTGAATATCTATCTCTCTCTTTT 1815  
Qy 661 caatatattgaagcattatcaggggtatgtctatgtgagggataacataatttgatgt 720  
Db 1816 CAATATTATTGAACATTATACAGGGTTATGTCTCATGTAGCGGAATACATATTGATGT 1875  
Qy 721 atttgaataaataacaataaggggttcgcgcgcacattcccccgaagaatgcacactgac 780

Db 1876 ATTTGAAAATAAATAAGGCTTCCGCGACATTCCCGAAGTCCACCTGAC 1935  
Qy 781 gtctagaacacattatcatgacattgaactaaatagcgtatcacgagcc 840  
Db 1936 GTCTAGAAACCATTAATATCATGACATTAACCTATAAATAGCGGTATACGAGGCC 1995  
Qy 841 ttctgtctgcgcgtttcgtgtgtaagggtaaaaacccctgacacatgcaagctccgag 900  
Db 1996 TTTCCCTCGCGCGCTTTGCGTGATGACGGTGAAGAAACCTGTGACATGACCTCCGAG 2055  
Qy 901 acggtcacagctgtctgttaagcgatgccggagacagacccgtcaaggcgctca 960  
Db 2056 ACGGTACAGCTGTGTGTAAAGCGGATCCGGAGACAGCAAGCCCGTACAGGCGCGTCA 2115  
Qy 961 gcgggtgttgcgaggtgctgcgggctgcttaactatgcgc 1001  
Db 2116 GCGGCTGTGGGGGTGTGCGGCGCTTGAACATATGCGGC 2156

RESULT 14  
US-09-039-982A-30/C  
Sequence 30, Application US/0903982A  
Patent No. 6225042  
GENERAL INFORMATION:  
APPLICANT: Cal, Zelling  
APPLICANT: Sprent, Jonathan  
APPLICANT: Brunmark, Anders  
APPLICANT: Jackson, Michael  
APPLICANT: Peterson, Per A  
TITLE OF INVENTION: ANTIGEN PRESENTING SYSTEM AND METHODS FOR ACTIVATION OF T-CELL  
NUMBER OF SEQUENCES: 59  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Olson & Hieltz, Ltd.  
STREET: 20 No. 6225042th. Wacker Drive, Suite 3000  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/039, 982A  
FILING DATE: 16-MAR-1998  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Olson, Arne M.  
REGISTRATION NUMBER: 30,203  
REFERENCE/DOCKET NUMBER: TSRI4710  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 580-1180  
TELEFAX: (312) 580-1189  
INFORMATION FOR SEQ ID NO: 30:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3883 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-09-039-982A-30

Query Match 100.0%; Score 1001; DB 4; Length 3883;  
Best Local Similarity 100.0%; Pred. No. 2.4e-301;  
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 cctgaacttatccgcctcatcagcttaattgttgcgggagcctcaagatgaagt 60  
Db 2499 CCTGCACTTATCCGCTCCATCCAGTCTATTAATTGTTCGCGGAAGCTAGAGTAA 2440

Qy 61 agtgcgcagttatagtttgcgcaacgtgtgtccattgtctaacagcagcgtgtgtca 120  
Db 2439 AGTCCGCACTTATATATTGGCGAAGCTGTGGCATCTCTCAAGGACATCGTGCTCA 2380  
Qy 121 cgcctgtcgtttgtatgctcatcattcaagctccggttcccaagatcaagcgattaca 180  
Db 2379 CGCTCGCTGTTGGTATGAGCTTATTCAGCTCCGGTCCCAAGATCAAGGAGTTTACA 2320  
Qy 181 tgaatccccatltttgtcaaaaaagcggttagtctcttgccttcgtctcgtgtcaga 240  
Db 2319 TGATCCCCCATGTTGTCAAAAAAGCGGTTAGCTCTTCGCTCCGATCGTTGTCA 2260  
Qy 241 agtaagtgtgcgcgaggttatcatcattatgttatgcaaacactgtacatacttctact 300  
Db 2259 AGTAAGTTGGCCGACAGTGTATCATCATATGTTATGACACATGATATTTCTTACT 2200  
Qy 301 gtcattgcacatcgttaagaagcttcttctgtgactgtgtgactcaaaccaagtcattcga 360  
Db 2199 GTCATGCCATCCGTAGATGCTTTCTGTGACTGTGAGTACTCAACCAAGTCATTTCTGA 2140  
Qy 361 gaatagtatagcgagcagcagttgtcttgcggcggtcaatacggataataaccgag 420  
Db 2139 GAATAGTGTATGCGGCGACCGAGCTTGCTTGCCTCCGCTCAATACGGGATATATACCGC 2080  
Qy 421 ccaatagcagaacttaaaagtgtcatcatcatatgtaaaaagcttctcggcgcaaaaac 480  
Db 2079 CCACATGACGAACCTTAAAGTGTATCATATGTAAGAAAGTCTTCCGGGCGCAAAAC 2020  
Qy 481 tcaaggaacttaccgctgttgaatcagtcagtcagtcagtcagtcagtcagtcagtcag 540  
Db 2019 TCAAGGATCTTACCGCTGTGAGATCCAGTTCAGTATCAACCTCGTGCACCACTCA 1960  
Qy 541 tcttcaacatcttcttcttccacagcgttctggttgagcaaaaacaggaagcaaat 600  
Db 1959 TCTTCACATCTTTTACTTTCACACAGGTTTCTGGGTGACAAACAGAGCAAAAT 1900  
Qy 601 gccgcaaaaaaggaataaaggcgacagcaagaaatgttgaactactacttcttctt 660  
Db 1899 GCCGCAAAAAGGAATTAAGGCGACACGGAATGTTGATATCTACTCTCTTTT 1840  
Qy 661 caatatattgaagcattatcaggggttattgtctcatagagcgatataattgaagt 720  
Db 1839 CAATATATTGAAGCATTTTATCAGGGTTATGTCTCATGAGCGGATATTTGAATGT 1780  
Qy 721 attagaataataaacaataggggttccgagcattcccgaaagtccactgac 780  
Db 1779 ATTTAGAAAATTAACAAATAGGGGTTCCGCGACATTTCCCGAAAGTCCACCTGAC 1720  
Qy 781 gtctagaacacattatcatgacattgaactaaatagcgtatcacgagcc 840  
Db 1719 GTCTAGAAACCATTAATATCATGACATTAACCTATAAATAGCGGTATACGAGGCC 1660  
Qy 841 ttctgtctgcgcgtttcgtgtgtaagggtaaaaacccctgacacatgcaagctccgag 900  
Db 1659 TTTCCCTCGCGCGCTTTGCGTGATGACGGTGAAGAAACCTGTGACATGACCTCCGAG 1600  
Qy 901 acggtcacagctgtctgttaagcgatgccggagacagacccgtcaaggcgctca 960  
Db 1599 ACGGTACAGCTGTGTGTAAAGCGGATCCGGAGACAGCAAGCCCGTACAGGCGCGTCA 1540  
Qy 961 gcgggtgttgcgaggtgctgcgggctgcttaactatgcgc 1001  
Db 1539 GCGGCTGTGGGGGTGTGCGGCGCTTGAACATATGCGGC 1499

RESULT 15  
US-09-641-30/C  
Sequence 30, Application US/09039641  
Patent No. 6251627  
GENERAL INFORMATION:  
APPLICANT: Cal, Zelling  
APPLICANT: Sprent, Jonathan

```

APPLICANT: Brunmark, Anders
APPLICANT: Jackson, Michael
APPLICANT: Peterson, Per A
TITLE OF INVENTION: ANTIGEN PRESENTING SYSTEM AND METHODS FOR
TITLE OF INVENTION: ACTIVATION OF T-CELLS
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESS: Olson & Hierl, Ltd.
STREET: 20 No. 6251627th Wacker Drive, Suite 3000
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/039,641
FILING DATE: 8-MAR-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Olson, Arne M.
REGISTRATION NUMBER: 30,203
REFERENCE/DOCKET NUMBER: TSRI4710
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 580-1180
TELEFAX: (312) 580-1189
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 3883 base pairs.
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOHETICAL: NO
ANTI-SENSE: NO
US-09-039-641-30

```

```

Query Match 100.0%; Score 1001; DB 4; Length 3883;
Best Local Similarity 100.0%; Pred. No. 2,4e-301;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 cctgcaacttaccgctccatccatccagctcattatggttcgaggaagctagaagtaagt 60
DB 2499 CCTGCACTTTATCCGCTCCATCCAGCTTATTAATTGTTGCCGGAGCTAGAGTAAGT 2440
QY 61 agtgcgcaatlaaagttgacgaacgltgtgcattgtctacaggaacgltgtgtca 120
DB 2439 AGTTCCGCACTTATAGTTTGGCAAGCTGTGGCCATTGCTACAGGCAATCGTGTGCA 2380
QY 121 cgtctcgtcttgatgagctcattcaagctcgggttcccaagatcaagaggaagttaca 180
DB 2379 CGCTCGCTGTTTGGTATGAGGCTTCAATCCGCTCCGTTCCCAAGATCAAGCGAGTTACA 2320
QY 181 tgatcccccatgtgtgcaaaaagcggttagctccctcggtccctccgactcgtgtcaga 240
DB 2319 TGATCCCCCATGTTGTGCAAAAAGCGGTTAGCTCTTCGCTCCGATCGTTGTGACA 2260
QY 241 agtaagttgacgagtgatatacactcaatggtatgtgcaagcaactgataatctctact 300
DB 2259 AGTAAGTTGGCCGACGAGTATCATCATGTTATGACGACACTGCATATATCTCTTACT 2200
QY 301 gtaatgcacatcgttaagaatgtcttctgtgactggtgagtaactcaaccaagtaattctga 360
DB 2199 GTCATGCCATCCGTAAGATCTTTTGTGACTGTGAGTACACCAAGTCAATCTCTGA 2140
QY 361 gaatagtgatgacgagcaacgaggtgtctgtgcccggcgctcaatacaggaataaccgag 420
DB 2139 GAATAGTGATGCGGCGACCGAGTGTCTTGGCCCGGCTCATACGGAGTAATACCGCG 2080

```

```

QY 421 ccacataagcaacttlaaagtgctcaatctggaagaacgtttctcgggacggaacatc 480
DB 2079 CCACATACAGCACTTTAAAGTGCTCATCATGTGAAAACGTTCTCGGGGCAAAACTC 2020
QY 481 tcaagatccttaacgctgttgaatccagttcagatgtaacccatcgtgtcaacccactga 540
DB 2019 TCAAGATCTTACCGCTGTTGATCCAGTCCGATGTAACCCACTCGTGCACCCACTGA 1960
QY 541 tcttaagatccttacttaccacagcgtttctgtgtgagcaaaaacggaagcaaat 600
DB 1959 TCTTAGCATCTTTACTTTTACACGCGTTTCTGGGTAGCAAAAACAGCAAGCAAAAT 1900
QY 601 gccgcaaaaaggaataaagggcagacggaatgltgaatacactactctctctt 660
DB 1899 GCCGCAAAAAGGAATTAAGGCGCACAGCAAAATGTTGAATACATCATCTTCTTTT 1840
QY 661 caattatgaagcattatcaaggttatgtgtcctcagagcggaatcattgaatgt 720
DB 1839 CAATATTTATGAAAGCATTTATCAGGGTTATGTCTCATGACCGGATACATATTTGAATGT. 1780
QY 721 attagaataataacaataaggggttcgcgcacattcccgaaaagtgccactgac 780
DB 1779 ATTTGAAAATAATAACAATAGGGGTTCCGCGCACATTTCCCGCAAAAGTCCACCTGAC 1720
QY 781 gctaaagaacacattatatacatgacataaactataaaaatagggcgtatcaagagccc 840
DB 1719 GCTTAAGAAACCATTAATATCATGACATTAACCTATAAAATAGGCGGTATCAGAGGCC 1660
QY 841 ttgtgtcgcggttctgtgtgagcgtgtaaaaccttcaacatgacagctccggag 900
DB 1659 TTTGCTCGCGCGTTTGTGATGACGCTAANAACCTCTACACATGACAGCTCCGGAG 1600
QY 901 acggtcacagctgtctgtgaagcgatgcggaagcagaacagccgtcagggcgctca 960
DB 1599 ACGGTACAGCTTGTCTGTAGCGGATGCCGGAGACACAAAGCCGTCAGAGCGCGTCA 1540
QY 961 gcgggtgttgcggggtgtcgggggtcgttaactatcgctgcgc 1001
DB 1539 GCGGCTGTGGCGGGGTGTGCGGCTGCTTAACATATGCGGC 1499

```

```

Search completed: January 8, 2002, 17:20:23
Job time: 12274 sec

```

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd

OM nucleic - nucleic search, using sw model

Run on: January 8, 2002, 17:22:58 ; Search time 7258.17 seconds (without all comments)

2272.912 Million cell updates/sec

```
Title: US-09-740-211-14_COPY_1_1000
Perfect score: 1000
```

Sequence: 1 cgccctgcagcagctgcg.....cagcctcatctgagcccta 1000

Scoring table:	IDENTITY_NUC	Correct	100
	Correct	100	100

Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum, Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

1: gb\_da: \*  
2: gb\_hlg: \*  
3: gb\_in: \*  
4: gb\_ov: \*  
5: gb\_om: \*  
6: gb\_pat: \*  
7: gb\_ph: \*  
8: gb\_pl: \*  
9: gb\_pro: \*  
10: gb\_ro: \*  
11: gb\_sts: \*  
12: gb\_sy: \*  
13: gb\_un: \*  
14: gb\_vi: \*  
15: em\_ba: \*  
16: em\_fun: \*  
17: em\_hum: \*  
18: em\_in: \*  
19: em\_om: \*  
20: em\_or: \*  
21: em\_ov: \*  
22: em\_pat: \*  
23: em\_ph: \*  
24: em\_pl: \*  
25: em\_ro: \*  
26: em\_sts: \*  
27: em\_sy: \*  
28: em\_un: \*  
29: em\_vi: \*  
30: em\_hlg\_hum: \*  
31: em\_hlg\_inv: \*  
32: em\_hlg\_rod: \*  
33: em\_hlg\_hum: \*  
34: em\_hlg\_inv: \*  
35: em\_hlg\_rod: \*  
36: em\_hlg\_other: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1000	100.0	4999	6	AR18378	AR18378 Sequence
2	1000	100.0	4999	6	AR146888	AR146888 Sequence
3	736.2	73.6	11933	6	AR138377	AR138377 Sequence
4	736.2	73.6	11933	6	AR146887	AR146887 Sequence
5	592.2	59.2	5035	6	AR034084	AR034084 Sequence
6	592.2	59.2	5035	6	AR071306	AR071306 Sequence
7	592.2	59.2	5035	6	AR112722	AR112722 Sequence
8	590.6	59.1	8831	6	EO0527	EO0527 Complete CD
9	590.6	59.1	8967	6	AX052730	AX052730 Sequence
10	590.6	59.1	8967	6	I71409	I71409 Sequence 1
11	590.6	59.1	8967	6	HSFV11R	X01179 Human mRNA
12	590.6	59.1	9029	6	HUMFV11C	M14113 Human coagu
13	589	58.9	9009	6	AR003710	AR003710 Sequence
14	589	58.9	9009	6	AR029065	AR029065 Sequence
15	589	58.9	9009	6	AR126884	AR126884 Sequence
16	589	58.9	9009	6	I31901	I31901 Sequence 3
17	589	58.9	9009	6	I63424	I63424 Sequence 1
18	589	58.9	9009	6	HUMFV11I	K01740 Human coagu
19	586	58.6	8241	6	A05328	A05328 Synthetic F
20	586	58.6	8241	6	A07042	A07042 Artificial
21	585.8	58.6	7440	6	I08345	I08345 Sequence 2
22	585	58.5	7272	6	I05404	I05404 Sequence 3
23	583.4	58.3	7056	6	I27063	I27063 Sequence 1
24	579.2	57.9	9354	6	AR003585	AR003585 Sequence
25	577.8	57.8	7440	6	I08457	I08457 Sequence 1
26	528	52.8	4278	6	I08644	I08644 Sequence 4
27	528	52.8	4281	6	I08643	I08643 Sequence 2
28	528	52.8	4548	6	I08642	I08642 Sequence 3
29	528	52.8	4551	6	I08641	I08641 Sequence 1
30	528	52.8	4670	6	AR110040	AR110040 Sequence 1
31	528	52.8	6999	6	I77105	I77105 Sequence 1
32	513.8	51.4	1993	6	I02047	I02047 Sequence 1
33	513.8	51.4	1993	6	I02047	I02047 Sequence 1
34	452.6	45.4	7032	4	AF016234	AF016234 Canis fam
35	451.4	45.3	7145	4	AF049489	AF049489 Canis fam
36	451.4	45.1	7493	6	AR003712	AR003712 Sequence
37	451.4	45.1	7493	6	AR029067	AR029067 Sequence
38	451.4	45.1	7493	6	AR126886	AR126886 Sequence
39	451.4	45.1	7493	6	I63427	I63427 Sequence 7
40	424.8	42.5	6539	4	MUSCVC11I	MUSCVC11I
41	423.8	42.4	4334	6	SS049517	SS049517 Sus domest
42	423.8	42.4	4334	6	AR029098	AR029098 Sequence
43	423.8	42.4	4334	6	AR126917	AR126917 Sequence
44	423.8	42.4	6402	6	AR029097	AR029097 Sequence
45	360.2	38.0	553	4	AF160523	AF160523 Ovis arie

RESULT	1				
ARI38378					
LOCUS	ARI38378	4999 bp	DNA	PAT	16-JUN-2001
DEFINITION	Sequence	14 from patent US 6200560.			
ACCESSION	ARI38378				
VERSION	ARI38378.1	GI:14480723			
KEYWORDS	.				
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	Unclassified.				
AUTHORS	1 (bases 1 to 4999)				
TITLE	Couto,L.B., Colosi,P.C. and Qian,X.				
JOURNAL	Adeno-associated virus vectors for expression of factor VIII by target cells				
FEATURES	Patent: US 6200560-A 14 13-MAR-2001;				
SOURCE	Location/Qualifiers				
	1..4999				
	/organism="unknown"				
BASE COUNT	1375 a	1151 c	1153 g	1318 t	
ORIGIN					





[illegible]

QY	351	tcttaataacatctactgacgaacgaatcaactttctttctctcacaggatcga-	409
Db	361	TCGCTGCCCTTGAATTAATCTGACACTGACATCCACTTTTCTTTTCTCCACAGGATGAT	420
QY	410	tccaacatgacaaatagagctctccacctgctctcttctgctctttgcgattcgtct	469
Db	421	TTCCACATGCAAAATAGAGCTCTCCACCGCTCTTCTTCTGTGCTTTTGCGATTCGCTTT	480
QY	470	agtgccaccagaagaatactactcctggtgtgcagtggaactgtcatabggagctatabgcaag	529
Db	481	AGTGCCACCAAGAAATACTACTCCTGGGTGCACATGGAACGTGTCATGCGACTATATGCAAGT	540
QY	530	gactcggtagagcgctgcctgtgtgacgcaagaattctctccaaagtgccaaaatctttcca	589
Db	541	GATCTCGGTGTGAGCTGCTCTGTGAGCCAGATTTTCTCTCTTAAGTGGCCAAATCTTTTCCA	600
QY	590	ttcaacacctcagtcgtgtgtacaaaagaactcgtttgtagaatccaagatccacttttc	649
Db	601	TTCAACACCTCAGTACGTCGTACAAAAGACTCTGTTGTAGAAATTCACGAGATCACCTTTTC	660
QY	650	aacatcgctaaagccaaggccacacttgaatgggtctgtcaggtctctacatccaaagctgag	709
Db	661	AACATCGCTAAGCCCAAGGCCACCCCTGGATGGGTGCTGATGGTCTTACCATCATCAAGCTAG	720
QY	710	gtttatgtctacagtgatgtcattacacttaagaacatgagcttccalcctgtcagcttcat	769
Db	721	GTTTATGTATACAGTGTGATTAATTAACACTTAAAGAACATGAGCTTCCACATCCTGATGCTTCAT	780
QY	770	gctgtgtgtatctactactggaagctctctgaaggagctgtgaatagatgacagaccagt	829
Db	781	GCTGTGTGTGTATCTACTGTGGAACCTTCTGTGAGAGCTGATATAGATGATCAGCCAGT	840
QY	830	caaaaggagaaagaaatgataaagtccttcctctgtgtgaagccatataatgtctgcag	889
Db	841	CAAAAGGAGAAAGAAAGATGATTAAGTCTTCCCTGTGTGGAAGCCATACATATGTCTGGAG	900
QY	890	gtccctgaagaagaatggttccaaatgagcctctgaccccaactgtgctctactacatatactt	949
Db	901	GTCCGTAAGAAGAAATGATGTCATGAGCTGTGACCCATGTGCTTACCTACATATATCTT	960
QY	950	tctcatgtgagacctgtgtaaagaacttgaatttaaggccctcatgttgagggcccta	1000
Db	961	TCTCATGTGTGACCTGTGTAAGAAGCTTGAATTCAGGCTTCATTTGAGAGCCCTA	1011
RESULT	4		
LOCUS	ARI46887	11933 bp	DNA
DEFINITION	Sequence 13 from patent US 6221349.		PAT
ACCESSION	ARI46887		
VERSION	ARI46887.1	GI:15110690	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 11933)		
AUTHORS	Conto,L.B., Colosi,P.C. and Qian,X.		
TITLE	Adeno-associated vectors for expression of factor VIII by target cells		
JOURNAL	Patent: US 6221349-A 13 24 APR-2001;		
FEATURES	Location/Qualifiers		
source	1..11933		
BASE COUNT	3258 a 2818 c 2717 g 3140 t		
ORIGIN	/organism="unknown"		
Query Match	73.6%;	Score 736.2;	DB 6;
Best Local Similarity	86.1%;	Pred. No. 4,4e-195;	Length 11933;
Matches 870;	Conservative 0;	Mismatches 118;	Indels 23;
			Gaps 4.
QY	13	caagctgcgctgcctgcctcactatgagggccggccgaagccgggagctgcggcgacc	72



ORGANISM Unknown.  
REFERENCE Unclassified.  
AUTHORS 1 (bases 1 to 5035)  
TITLE Hybrid proteins with modified activity  
JOURNAL Patent: US 5910481-A 1 08-JUN-1999;  
FEATURES Location/Qualifiers  
source 1..5035  
/organism="unknown"

BASE COUNT 1484 a 1127 c 1110 g 1314 t  
ORIGIN

Query Match 59.2%; Score 592.2; DB 6; Length 5035;  
Best Local Similarity 97.9%; Pred. No. 8.9e-155;  
Matches 600; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

OY tcttttccacaggtatcgatccacatgcaatagaagctctccactgtcttct 447  
DB 7 TCCAGTTGAACATTTGTAGCAAGCCACCATGGAATAGAGCTCTCACCTCTTCTTCT 66  
OY gtgcctttgcatctgctttagtgccacagaaatgactctggtgagtggaact 507  
DB 448 gtgcctttgcatctgctttagtgccacagaaatgactctggtgagtggaact 507  
OY gtgcctttgcatctgctttagtgccacagaaatgactctggtgagtggaact 507  
DB 67 GTGCCTTTGGGATTTGCTTGTAGTCCACCAAGAAATGACTACCTGGGTGAGTGAAGT 126  
OY gtcatggagactatgcaaaagtatctggtgagtgctgctggtgagtggaact 567  
DB 127 GTCATGGGACTATATGCAAAAGTATCTGCTGAGTGGTGGAGCAAGATTTCTCTCC 186  
OY taaagtgccaaaatcttttccattcaaacctgagtcgtgtacaaaagaactctgtt 627  
DB 568 taaagtgccaaaatcttttccattcaaacctgagtcgtgtacaaaagaactctgtt 627  
OY taaagtgccaaaatcttttccattcaaacctgagtcgtgtacaaaagaactctgtt 627  
DB 187 TAGAGTGCCAAAATCTTTTCCATTCAACACTGAGTGTGACAAAAGACCTGTTTGT 246  
OY agaattcaggaatcacttttcaacatcgctaaagagccacccctgagtggtctgct 687  
DB 628 agaattcaggaatcacttttcaacatcgctaaagagccacccctgagtggtctgct 687  
OY aggtctccatccagctggtgagtggttataatagatgagtgatcttaagaacatgagc 747  
DB 247 AGAATTCACGAGTACCTTTTCAACATCTGTAAGCCAAAGGCCACCTGGATGGTCTGCT 306  
OY aggtctccatccagctggtgagtggttataatagatgagtgatcttaagaacatgagc 747  
DB 307 AGGTCTCACATCCAGCTGAGGTTTATGATGATGATGATGATGATGATGATGATGATG 366  
OY tttccatccatccagctggtgagtggttataatagatgagtgatcttaagaacatgagc 807  
DB 748 tttccatccatccagctggtgagtggttataatagatgagtgatcttaagaacatgagc 807  
OY tttccatccatccagctggtgagtggttataatagatgagtgatcttaagaacatgagc 807  
DB 367 TTTCCATCCCTGTCAGTCTTCTGATCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 426  
OY tgaatatgatatgatacgaagcaagcaagcaagcaagcaagcaagcaagcaagcaagcaag 867  
DB 808 tgaatatgatatgatacgaagcaagcaagcaagcaagcaagcaagcaagcaagcaagcaag 867  
OY tgaatatgatatgatacgaagcaagcaagcaagcaagcaagcaagcaagcaagcaagcaag 867  
DB 427 TGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 486  
OY aagccatataatgctgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 927  
DB 868 aagccatataatgctgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 927  
OY aagccatataatgctgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 927  
DB 487 AAGCCATACATATGCTGAGAGGCTCTGGAAGAAAGATGATGATGATGATGATGATGATGATG 546  
OY gtgccttactactacatacttctcattgtagcctggtgaaagacttgaattcagagcct 987  
DB 928 gtgccttactactacatacttctcattgtagcctggtgaaagacttgaattcagagcct 987  
OY gtgccttactactacatacttctcattgtagcctggtgaaagacttgaattcagagcct 987  
DB 547 GTGCCTTACTACTACTATCTTCTCATGTGACCTGTGTAAGAAAGCTTGAATTCAGGCT 606  
OY catggagagcccta 1000  
DB 988 catggagagcccta 1000  
OY catggagagcccta 1000  
DB 607 CATTGAGGCCCTA 619

RESULT 7  
LOCUS AR112722 5035 bp DNA--  
DEFINITION Sequence 1 from patent US 6130203.  
ACCESSION AR112722  
VERSION AR112722.1 GI:14092622  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 5035)

AUTHORS Voorberg,J.J.  
TITLE Hybrid proteins with modified activity  
JOURNAL Patent: US 6130203-A 1 10-OCT-2000;  
FEATURES Location/Qualifiers  
source 1..5035  
/organism="unknown"

BASE COUNT 1484 a 1127 c 1110 g 1314 t  
ORIGIN

Query Match 59.2%; Score 592.2; DB 6; Length 5035;  
Best Local Similarity 97.9%; Pred. No. 8.9e-155;  
Matches 600; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

OY tcttttccacaggtatcgatccacatgcaatagaagctctccactgtcttct 447  
DB 7 TCCAGTTGAACATTTGTAGCAAGCCACCATGGAATAGAGCTCTCACCTCTTCTTCT 66  
OY gtgcctttgcatctgctttagtgccacagaaatgactctggtgagtggaact 507  
DB 448 gtgcctttgcatctgctttagtgccacagaaatgactctggtgagtggaact 507  
OY gtgcctttgcatctgctttagtgccacagaaatgactctggtgagtggaact 507  
DB 67 GTGCCTTTGGGATTTGCTTGTAGTCCACCAAGAAATGACTACCTGGGTGAGTGAAGT 126  
OY gtcatggagactatgcaaaagtatctggtgagtgctgctggtgagtggaact 567  
DB 127 GTCATGGGACTATATGCAAAAGTATCTGCTGAGTGGTGGAGCAAGATTTCTCTCC 186  
OY taaagtgccaaaatcttttccattcaaacctgagtcgtgtacaaaagaactctgtt 627  
DB 568 taaagtgccaaaatcttttccattcaaacctgagtcgtgtacaaaagaactctgtt 627  
OY taaagtgccaaaatcttttccattcaaacctgagtcgtgtacaaaagaactctgtt 627  
DB 187 TAGAGTGCCAAAATCTTTTCCATTCAACACTGAGTGTGACAAAAGACCTGTTTGT 246  
OY agaattcaggaatcacttttcaacatcgctaaagagccacccctgagtggtctgct 687  
DB 628 agaattcaggaatcacttttcaacatcgctaaagagccacccctgagtggtctgct 687  
OY aggtctccatccagctggtgagtggttataatagatgagtgatcttaagaacatgagc 747  
DB 247 AGAATTCACGAGTACCTTTTCAACATCTGTAAGCCAAAGGCCACCTGGATGGTCTGCT 306  
OY aggtctccatccagctggtgagtggttataatagatgagtgatcttaagaacatgagc 747  
DB 307 AGGTCTCACATCCAGCTGAGGTTTATGATGATGATGATGATGATGATGATGATGATG 366  
OY tttccatccatccagctggtgagtggttataatagatgagtgatcttaagaacatgagc 807  
DB 748 tttccatccatccagctggtgagtggttataatagatgagtgatcttaagaacatgagc 807  
OY tttccatccatccagctggtgagtggttataatagatgagtgatcttaagaacatgagc 807  
DB 367 TTTCCATCCCTGTCAGTCTTCTGATCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 426  
OY tgaatatgatatgatacgaagcaagcaagcaagcaagcaagcaagcaagcaagcaagcaag 867  
DB 808 tgaatatgatatgatacgaagcaagcaagcaagcaagcaagcaagcaagcaagcaagcaag 867  
OY tgaatatgatatgatacgaagcaagcaagcaagcaagcaagcaagcaagcaagcaagcaag 867  
DB 427 TGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 486  
OY aagccatataatgctgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 927  
DB 868 aagccatataatgctgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 927  
OY aagccatataatgctgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 927  
DB 487 AAGCCATACATATGCTGAGAGGCTCTGGAAGAAAGATGATGATGATGATGATGATGATGATG 546  
OY gtgccttactactacatacttctcattgtagcctggtgaaagacttgaattcagagcct 987  
DB 928 gtgccttactactacatacttctcattgtagcctggtgaaagacttgaattcagagcct 987  
OY gtgccttactactacatacttctcattgtagcctggtgaaagacttgaattcagagcct 987  
DB 547 GTGCCTTACTACTACTATCTTCTCATGTGACCTGTGTAAGAAAGCTTGAATTCAGGCT 606  
OY catggagagcccta 1000  
DB 988 catggagagcccta 1000  
OY catggagagcccta 1000  
DB 607 CATTGAGGCCCTA 619

RESULT 8  
LOCUS E00527 8831 bp RNA  
DEFINITION Complete cDNA encoding human factor VIII.  
ACCESSION E00527  
VERSION E00527.1 GI:2168806  
KEYWORDS JP 1985243023-A/2.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens.  
REFERENCE 1 (bases 1 to 8831)  
AUTHORS Danilov,J.K., Richiyado,M.R., Goodon,A.B. and Uritanu,A.U.  
TITLE MANUFACTURE OF FUNCTIONAL VIII FACTOR



APTKDEDCAMAFVSVDLDEKDVHSLIGLPLVCHTNTLNPAGROYVDFALFET  
 IFEDEKSWYFTEEMERNCRAPCNIOMEPTREKREAFINGIMDLPGVLMQDOR  
 IRWLLSMGSENINISIFSGHVFREKREKALNYLYPVEVEEMLEPKAGIWR  
 VECLIGELHAGMSHFLVYNSKOCOTPLGMSAGHLPDOITASOYGOVMPKRLIYR  
 SSTINAMSTEKEPFIKIDLLAPMIIEIKTQAGROKRSLSLYISOFIYMSLDKKNM  
 TYRGNSTETLWFEENVDSSGKIKHNPPLIARIRLPHYSIRSLRMEIAGCDL  
 NSCMPLEMSKALSDAQITASSFTNPFATWSRKALHLOGSNMARPVNNPKEX  
 LOVDFOKTMKVTGKVSGLTSMYKEFLISSODGHOWITLFEONKVKVFOGNO  
 DSFTPVNSLDPPLLTRILYRIHPOSWHQAIALRMEVLCECAODLY"

mat. peptide 167. 7162  
 BASE COUNT 2841 a 1898 c 1833 g 2395 t  
 ORIGIN

Query Match 59.1%; Score 590.6; DB 6; Length 8967;  
 Best Local Similarity 96.2%; Pred. No. 2.5e-154;  
 Matches 605; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 372 actgaacatccactttctttcttccacagatcgatccacacgaagaagactct 431  
 DB 66 ATTACCTTTTGGCTTCCAGTTGAACATTTGTAGCAATAGTCAGCAATAGAGCTCT 125  
 QY 432 ccacatgctcttcttctgctcttctgcatctgctttagtgccacagaagatactacc 491  
 DB 126 CCACCTGCTCTTCTTCTGCTCTTTGCGATTCTGCTTTAGTGCCACAGAAATACTACC 185  
 QY 492 tgggtgcagtggaactgcatggaactatatagcaagtgatctcggtgagctgctg 551  
 DB 186 TGGGTGAGTGAAGACTGCTATGGAAGTGTATGCGGTGAGTGGCTGCTG 245  
 QY 552 accgaagatctctctctagagtgcaaaatcttccatccacacactcgctgtata 611  
 DB 246 ACGAAGATTTCTCTCTAGAGTCCCAAAATCTTTTCATTCAACACTCAGTGTGTACA 305  
 QY 612 aaaaagactctgtttagaattacagatcacctttcaacatcgctgaagccac 671  
 DB 306 AAAAGACTCTGTTTGTAGATTCACGATCACCTTTCAACATCGCTTAAGCCAGGCCAC 365  
 QY 672 cctgagatgggtctgctagtgctctacacacagctgaggtttatgataagtgatata 731  
 DB 366 CCGTGAATGGGTCTGCTAGTGTCTACATCCAGCTGAGTTTATGATACAGTGTCTATTA 425  
 QY 732 caactaagaacatgcttccatccatctgctgcttctgctgttcttctactgta 791  
 DB 426 CACTTAAGAACATGCTTCCATCTCTCACTGCTTCACTGCTTGTGCTACTACTGGA 485  
 QY 792 aagcttctgagagctggaatatagatacagaccagtcagaagggagaagaagatgata 851  
 DB 486 AAGCTTCTGAGGAGCTGGAATATGATGATCAAGCAAGTCAAGGAGAAAGATGATA 545  
 QY 852 aagcttccctgtgtggaagccatataatgtctgagagctctggaagagaatggtccaa 911  
 DB 546 AAGCTTCTGCTGAGGAGCCATATATGCTGCGAGGCTCTGAAAGAAATGCTCAA 605  
 QY 912 tggcctctgaccacactgtgcttactactacatatacttctcaltgtygacctgtaaaag 971  
 DB 606 TGGCCTCTGACCCACTGTGCTTACTACTATCTTTTCATGATGAGACTGCTGTAAGAAG 665  
 QY 972 actgaatcagagcctcatggaagccta 1000  
 DB 666 ACTGAATTCAGGCTCATTTGAGCCCTTA 694

RESULT 10  
 LOCUS I71409 8967 bp DNA PAT 03-APR-1998  
 DEFINITION Sequence 1 from patent US 5681746.  
 ACCESSION I71409  
 VERSION I71409.1 GI:3007544  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 Unclassified

REFERENCE 1 (bases 1 to 8967)  
 AUTHORS Bodner, M., de Polo, N.J., Chang, S., Hsu, D. Chi-Tang and Respass, J. G.  
 TITLE Retroviral delivery of full length factor VIII  
 JOURNAL Patent: US 5681746-A 1 28-Oct-1997.  
 FEATURES location/Qualifiers  
 source 1..8967  
 /organism="unknown"

BASE COUNT 2841 a 1898 c 1833 g 2395 t  
 ORIGIN

Query Match 59.1%; Score 590.6; DB 6; Length 8967;  
 Best Local Similarity 96.2%; Pred. No. 2.5e-154;  
 Matches 605; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 372 actgaacatccactttctttcttccacagatcgatccacacgaagaagactct 431  
 DB 66 ATTACCTTTTGGCTTCCAGTTGAACATTTGTAGCAATAGTCAGCAATAGAGCTCT 125  
 QY 432 ccacatgctcttcttctgctcttctgcatctgctttagtgccacagaagatactacc 491  
 DB 126 CCACCTGCTCTTCTTCTGCTCTTTGCGATTCTGCTTTAGTGCCACAGAAATACTACC 185  
 QY 492 tgggtgcagtggaactgcatggaactatatagcaagtgatctcggtgagctgctg 551  
 DB 186 TGGGTGAGTGAAGACTGCTATGGAAGTGTATGCGGTGAGTGGCTGCTG 245  
 QY 552 accgaagatctctctctagagtgcaaaatcttccatccacacactcgctgtata 611  
 DB 246 ACGAAGATTTCTCTCTAGAGTCCCAAAATCTTTTCATTCAACACTCAGTGTGTACA 305  
 QY 612 aaaaagactctgtttagaattacagatcacctttcaacatcgctgaagccac 671  
 DB 306 AAAAGACTCTGTTTGTAGATTCACGATCACCTTTCAACATCGCTTAAGCCAGGCCAC 365  
 QY 672 cctgagatgggtctgctagtgctctacacacagctgaggtttatgataagtgatata 731  
 DB 366 CCGTGAATGGGTCTGCTAGTGTCTACATCCAGCTGAGTTTATGATACAGTGTCTATTA 425  
 QY 732 caactaagaacatgcttccatccatctgctgcttctgctgttcttctactgta 791  
 DB 426 CACTTAAGAACATGCTTCCATCTCTCACTGCTTCACTGCTTGTGCTACTACTGGA 485  
 QY 792 aagcttctgagagctggaatatagatacagaccagtcagaagggagaagaagatgata 851  
 DB 486 AAGCTTCTGAGGAGCTGGAATATGATGATCAAGCAAGTCAAGGAGAAAGATGATA 545  
 QY 852 aagcttccctgtgtggaagccatataatgtctgagagctctggaagagaatggtccaa 911  
 DB 546 AAGCTTCTGCTGAGGAGCCATATATGCTGCGAGGCTCTGAAAGAAATGCTCAA 605  
 QY 912 tggcctctgaccacactgtgcttactactacatatacttctcaltgtygacctgtaaaag 971  
 DB 606 TGGCCTCTGACCCACTGTGCTTACTACTATCTTTTCATGATGAGACTGCTGTAAGAAG 665  
 QY 972 actgaatcagagcctcatggaagccta 1000  
 DB 666 ACTGAATTCAGGCTCATTTGAGCCCTTA 694

RESULT 11  
 LOCUS HSEVIIIR 8967 bp mRNA PRI 21-MAR-1995  
 DEFINITION Human mRNA for factor VIII.  
 ACCESSION X01179  
 VERSION X01179.1 GI:31498  
 KEYWORDS factor VIII; signal peptide.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 REFERENCE 1 (bases 1 to 8967)  
 AUTHORS Wood, W. I., Capon, D. J., Simonsen, C. C., Eaton, D. L., Gitschler, J.,



**AUTHORS** Youssoufian, H., Wong, C., Aronis, S., Platakoukis, H., Kazazian, H. H., Jr. and Antonarakis, S. E.  
**TITLE** Moderately severe Hemophilia A resulting from Glu----Gly substitution in exon 7 of the factor VIII gene  
**JOURNAL MEDLINE** Am. J. Hum. Genet. 42 (6), 867-871 (1988)  
**REFERENCE** 88220354  
**AUTHORS** 3 (sites)  
 Bernardi, F., Legnani, C., Volinia, S., Patrascchini, P., Roderigo, G., DeRosa, V., and Marchetti, G.  
**TITLE** A HindIII RFLP and a gene lesion in the coagulation factor VIII gene  
**JOURNAL MEDLINE** Hum. Genet. 78 (4), 359-362 (1988)  
**COMMENT** [3] sites: mutation causing hemophilia.  
 [2] sites: mutations causing hemophilia.  
 Draft entry and Clean copy sequence for [1] kindly provided by M.A. Truett, 26-FEB-1986.  
 The mutation at position 1042 results in a change of Glu to Lys, and the one at position 1043 in a change from Glu to Gly. The mutation at position 1042 produces a premature stop codon.  
**FEATURES** Location/Qualifiers  
 source 1..9029  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /map="Xq28"  
 1..9029  
 /note="F.VIII mRNA"  
 172..7237  
 /gene="F8G"  
 172..7238  
 /gene="F8C"  
 /note="coagulation factor VIII:C signal peptide"  
 172..7227  
 /note="preprocoagulation factor VIII:C"  
 /gene="F8C"  
 /codon\_start=1  
 /db\_xref="GDB:600-119-124"  
 /protein\_id="AA52485.1"  
 /db\_xref="GI:182818"  
 /translation="MOELSTCFELCLRFCSATRRYVIGAVELSDMDSGLGELPVDAPRPVPSKSPENTSVYKTLFEFTDHLNIAKPRPMGLGPTLOAVYDTVITLKNASHVSLHAGVSVKMAEAEVDQDSQEKEDVFPDGSFTYVQVYLKENGPMASDPLCTGYSTLHYDIADYDINSLGALLVRESLSAKKMKOHLKFLFLAVDEGKSMHSTKSLMODDVAARPMVHNYNGVNSLGLCHCHKSYVMHYIGMTPEVHSLFEGHFLVNRQASLEISPTFLTAQLLMDLGFLEHCHSSHQHDGEAYVAVDSCPEPOLRMKNNEEADDDQLEMGVPRFDDNSQSLQIRSVAKKHPTVMHYIAEEDMDYAPLVLAADDRSYKQYELNNGVPRIGRYKRVPMAYTDEPKTREAIQHEGSLIGPLLYGEGVCTLLIFKNOSRPNYIPHGTIDVRLYSRIRPKGVKLKOPRLIGELPKFYKMTVYEDGPTKSDPCLTRVYSSPFNMRDASGLIGPLICTKESYDQNGNOMSDKRVLLFVFDENRMYLTENQRELPNAGVQLEDPEFQNSIMHSHNGVFDLSQVCLHVAWYLLISAGQDFLSVPSGTFKHKMYEDTLTLPFSGEIVFMSNMENGLMLICCHNSDENRMTALIKVSDCKMTAYEEDSYEDISAYLLSNNAIEPRFSQNSRHSTROKOFNATIPENDIEKTPFAHRTPMKIONVSSDLMLLROSPTPHGLSLDLOEKYETFDSDPSGALDNNLSLWHTFROLHSGDMVTPSPESGLQLRNEKLTGTAAETLKLDKFKVSSNNLISITPNDLAACDTNSSLGCPSPMVHYDSQDITLPEKKSPTLESQPSLSENNDSKLLSGLMNGQSSWKNVSTESGRLPFGKRAHGALLLTQDALRQVSI SLKTKNSLSSNTKRTHIDGPSLLIENSPVWONLIESDTEKRVYPLHDMMLDKNATALRLHMSNKTSSKNMENVQOKKEGIPDAQNPDMSPFKMLFLPSAIVIRTHKNSLNGOGSPKOLVSGPEKSVGONFLSEKNVYVVGKEGFEFDVGLKMPVSSNNLETLNLDNHNNTNHOEKLIOEIEKELIENNVLPQIHVYTGKMKMKNLPLSTRTONEGSDGAYAPVLQDRSLDSTNRTKRTAHFSKSGEENLFGCLNQKQLEYKACTRISPNTSOQNFVTSRSLKQFRLPLETELEKRIIVDDTSWKNMHLTPSTLQIDNYPEKEGATVQSPSLDCLTRSHSPIQANRSPDLAKVSSPFSIRIYLTFLFDNSHLPAAAYSRKDSGVQESSHFLQAGAKNNLSAILLENTGQRENGSLGATSNISYKVKENTVLPKDPVPSGVELLPKLVHYQDLPTETGQRENGSLGATSNISYKVKIKNENBRGKVPFLVATNESSAKTPSKLPLAMNHYGTOLPREKMSOKSPKTAERKKTIIISLANCESNHAIALINEGQNKPELIVYTAQGRERELCSQNPPLKRRQREITRTLLQSDQEEIDYDTISVEMKEDDIDEDNSQSPRSQKTRHYTAAVBRMLDYGMSSPHVLNRNAGSVPQPKRVVQDFGDFSTPLRGLNELHGLKGPVIAEVEDINMVFTRNAGSRPVSFSLISYEDDQAGAPKRNFKPNETKTYMKLVQNHMAATKDFQCKAAVPSVDYLEKDVSHGLPLVCHTUNLPAHGRQVTVQDFALFETIDETKSWTFETEMERNCRAPCNLOMEDPTFRENTRFALINQIMDTLPGLVMAQDQR

IRWYLIHSMGSENHSHIFHSQVFTVRKKEEYKMMALNLYPGVFETVEMLPKAGIMR  
 VECTIGEHNLHAGSGTLFLIYYSNKKDGHIRDPQITASGOGWAPKRLARHY  
 SCSINNAWSTKEPSPYIKVOLLAPMIITHGIKTGAGARPESSLYISQFIIMYSLDGKMO  
 TYGNSTGTLWVFNVDSSGIRHNIFNPITARIIRLPHYSTRSTIRMLKCDL  
 NSCSMPIGESKAI SDQITASSYFTNMTAPSLHLYGSRNAMPPOVNNREW  
 LQVDFQRTKRVGTVOGVKSLTSMYVKEFLISSQDQHOWTLFFQNKVAVFGNQ  
 DSFTPYVNSLDPELLRRLIRIHQSVWVHIALRMEVLCEADLY"  
 898..7224  
 /gene="F8C"  
 /note="coagulation factor VIII:C"  
 1042  
 /gene="F8C"  
 /note="g in wt; a in hemophiliac [2]"  
 1043  
 /gene="F8C"  
 /note="a in wt; g in hemophiliac [2]"  
 6853  
 /gene="F8C"  
 /note="c in wt; t in hemophiliac [3]"  
 BASE COUNT 2860 a 1910 c 1848 g 2411 t  
 ORIGIN  
 Query Match 59.1%: Score 590.6; DB 9; Length 9029;  
 Best Local Similarity 96.2%: Pred. No. 2,5e-154;  
 Matches 605; Conservative 0; Mismatches 24; Indels 0; Gaps 0;  
 372 actgacatccactttcttcttcttccacagagatcgatccacatgcaaaatagagctc 431  
 128 ATTAACCTTTGCTTCCACAGTGAACATTTGTAGCATTAAGTACATGCAATAGAGCTCT 187  
 432 ccactctcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 491  
 188 CCACCTGCTTCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTG 247  
 492 ttggttcagtggaactgcatgagacatataatgcaaaatgcatctgctgctgctg 551  
 248 TGGGTGACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 307  
 552 aagcaaatct 611  
 308 ACGCAAGATTCTTCCCTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 367  
 612 aaagactcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 671  
 368 AAAAGACTCTTGTGTAATTCAGGATCAGCTTTTCAACATCGCTTAAGCAAGGCCAC 427  
 672 cctgagatgct 731  
 428 CCGATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 487  
 732 caattaagaacatgctctccatctctctctctctctctctctctctctctctctctct 791  
 488 CACTTAAGAAATGAGCTTCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 547  
 792 aagcttcgaaggagctgaatataatgataatgataatgataatgataatgataatgata 851  
 548 AAGCTTCGAGAGGAGCTGAATATGATGATGATGATGATGATGATGATGATGATGATGAT 607  
 852 aagcttcct 911  
 608 AAGTTCCT 667  
 912 tggcctctgacccactgct 971  
 668 TGGCCTCTGACCCACATGCTGCTTACTCTACTATATCTTCTCATGTGAGCTGTGTAAG 727  
 972 actgaattcagagcctcatttgagcccta 1000  
 728 ACTGAATTCAAGGCTCAATTGAGGCCCTA 756

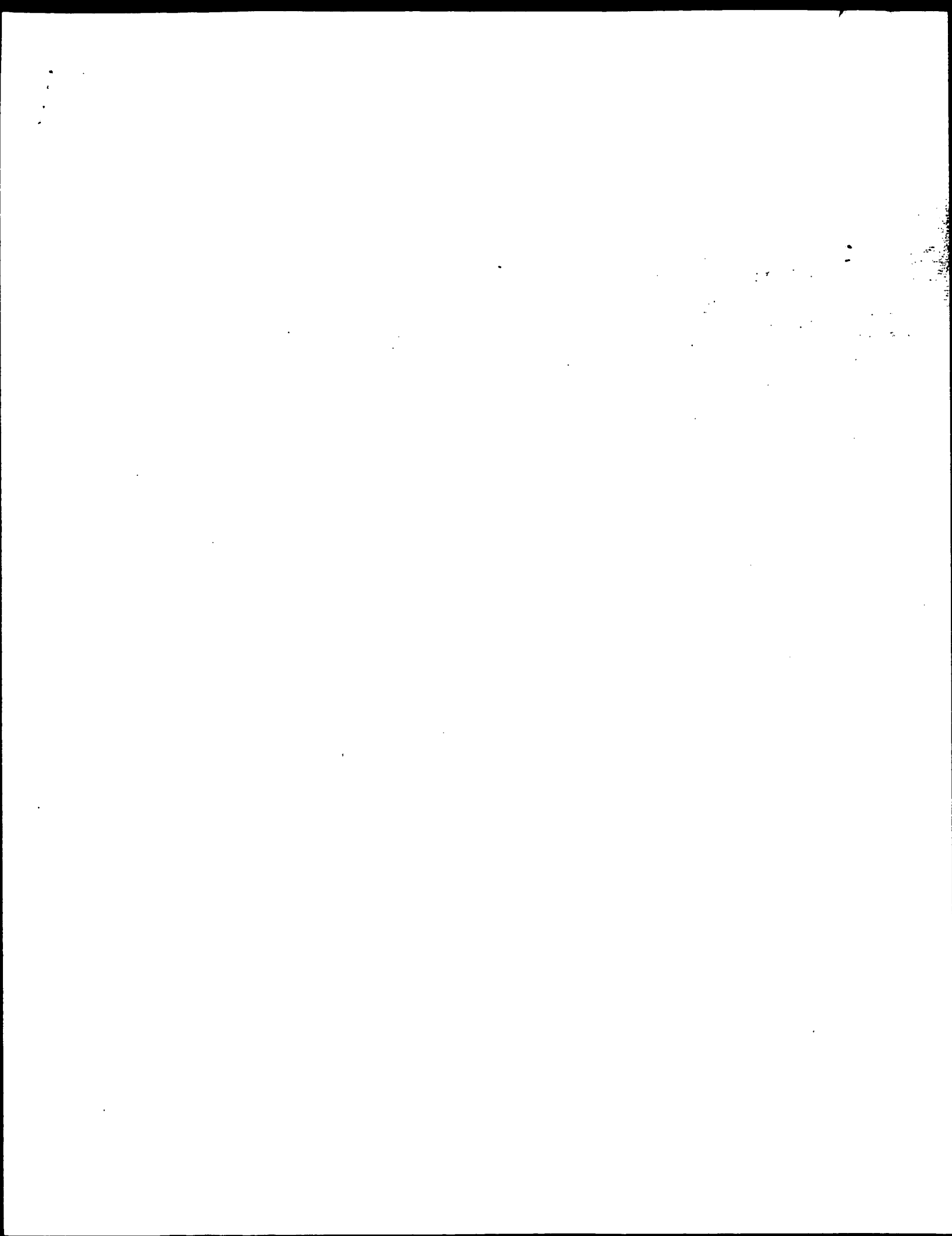
RESULT 13







SOURCE	ORGANISM	REFERENCE	TITLE	AUTHORS	BASE COUNT	ORIGIN
Unknown.	Unknown.	1 (bases 1 to 9009)	Modified Factor VIII	Lollar,J.S.	2853	a
		Unclassified.	Patent: US 6180371-A		1907	c.
			Location/Qualifiers		1844	g
			1..9009		2405	t
			/organism="unknown"			



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model.

Run on: January 8, 2002, 15:12:55 ; Search time 4541.38 Seconds  
(without alignments)  
2368.559 Million cell updates/sec

Title: US-09-740-211-13\_COPY\_6000\_7000

Perfect score: 1001

Sequence: 1 ccgcgaacttalcgcgcctc.....ggctggcttaactatgcgcgc 1001

Scoring table: IDENTITY\_NUC

Gapop 10.0, Gapext 1.0.

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: em\_estfun:\*  
2: em\_esthm:\*  
3: em\_estin:\*  
4: em\_estom:\*  
5: em\_estpl:\*  
6: em\_estda:\*  
7: em\_estro:\*  
8: em\_estov:\*  
9: em\_hic:\*  
10: gp\_est1:\*  
11: gp\_est2:\*  
12: gp\_hic:\*  
13: gp\_gss:\*  
14: em\_gss\_fun:\*  
15: em\_gss\_hum:\*  
16: em\_gss\_inv:\*  
17: em\_gss\_pln:\*  
18: em\_gss\_pro:\*  
19: em\_gss\_rtd:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	766.8	76.6	935	11	BG838279
2	741	74.0	841	10	AL042026
3	673.2	67.3	754	13	BG920379
4	650.6	65.0	767	10	AL040542
5	638	63.7	693	10	AL598527
6	636.6	63.6	688	10	AL042640
7	607.2	60.7	914	10	BE569261
8	591.4	59.1	702	10	AL042540
9	588.2	58.8	695	13	AG002183
10	576.8	57.6	658	13	AG014918
11	559.2	55.9	793	13	AG000364
12	554	55.3	696	13	AG009765

13	548.4	54.8	696	10	BE573230	BE573230 601333125
14	546.4	54.6	801	10	BE569687	BE569687 601332450
15	545.6	54.5	554	10	AL041808	AL041808 DKFZP434B
16	541.6	54.1	720	13	AG000534	AG000534 Homo sapi
17	531.8	53.1	711	13	AG010947	AG010947 Homo sapi
18	531.4	53.1	890	10	AL046135	AL046135 DKFZP434P
19	517.6	51.7	619	10	AV611386	AV611386 AV611386
20	508.6	50.8	710	13	CNS075M1	AL430631 clone BA0
21	499.6	49.9	575	13	AZ918759	AZ918759 1006006H0
22	496	49.6	620	10	AL599532	AL599532 DKFZP313D
23	488.6	48.8	604	13	AG002706	AG002706 Homo sapi
24	488	48.8	499	10	AV612734	AV612734 AV612734
25	476.6	47.6	551	10	AL041936	AL041936 DKFZP434L
26	467	46.7	683	11	BG838185	BG838185 GC01_06E0
27	466.2	46.6	890	13	B07159	B07159 CSRL72-cl-u
28	461	46.1	721	10	AL044583	AL044583 DKFZP434N
29	459.8	45.9	725	13	AG001126	AG001126 Homo sapi
30	457	45.7	647	13	AG008943	AG008943 Homo sapi
31	453	45.3	832	11	BG923768	BG923768 602825893
32	451.4	45.1	474	10	AV601733	AV601733 AV601733
33	451.2	45.1	684	13	AG014348	AG014348 Homo sapi
34	440	44.0	911	11	BG923556	BG923556 602823195
35	435	43.5	706	13	AG000533	AG000533 Homo sapi
36	433.2	43.3	686	13	AG016243	AG016243 Homo sapi
37	429	42.9	705	13	AG004170	AG004170 Homo sapi
38	429	42.9	705	13	AG010296	AG010296 Homo sapi
39	423	42.3	715	13	AG013813	AG013813 Homo sapi
40	418.8	41.8	751	13	AG007414	AG007414 Homo sapi
41	414	41.4	414	10	AV404393	AV404393 AV404393
42	413.6	41.3	501	10	AV596377	AV596377 AV596377
43	410	41.0	718	13	AG009622	AG009622 Homo sapi
44	409	40.9	640	13	B84895	B84895 RPI111-28L6
45	406	40.6	738	10	AL598454	AL598454 DKFZP313H

## ALIGNMENTS

RESULT 1	BG838279/c	935 bp	mRNA	EST	25-MAY-2001
LOCUS	GC01_10E07	R GC01_AAFc	ECORc_cold_stressed	Glycine-clandestina	
DEFINITION	Glycine clandestina cDNA clone GC01_10E07, mRNA sequence.				
ACCESSION	BG838279.1	GI:14204601			
VERSION	BG838279.1				
KEYWORDS	EST.				
SOURCE	Glycine clandestina.				
ORGANISM	Glycine clandestina				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;				
	Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;				
	Glycine.				
REFERENCE	1 (bases 1 to 935)				
AUTHORS	Singh, J.A., Farah, S., Chapados, J., Couroux, P., De Moors, A., Harris, L.J., Hattori, J.I., Ouellet, T., Robert, L.S., Spolt, D. and Tinker, N.A.				
TITLE	Expressed Sequence Tags from Cold-Stressed Glycine clandestina				
JOURNAL	Seedlings				
COMMENT	Unpublished (2001)				
	Contact: Singh, J.A.				
	Eastern Cereal and Oilseed Research Centre				
	Agriculture and Agri-food Canada				
	960 Carling Avenue, Bldg. 20, Ottawa, Ontario, K1A 0C6, Canada				
	Tel: (613) 759-1662				
	Fax: (613) 759-1701				
	Email: singhja@em.agr.ca.				
FEATURES	Location/Qualifiers				
source	1..935				
	/organism="Glycine clandestina"				
	/cultivar="1035"				
	/db_xref="taxon:45687"				
	/clone="GC01_10E07"				
	/clone_11b="GC01_AAFc_ECORc_cold_stressed_glycine-clandest				

ina  
/tissue\_type="Leaves, stem"  
/note="Vector:: Bluescript SK+/XhoI-EcoRI. Site\_1: EcoRI.  
Site\_2: XhoI. Plants incubated at 2 degrees under 12 hours  
of light/day. Harvested after only 2-3 days of cold  
treatment. cDNA was prepared with the Uni-Zap cDNA kit  
from Stratagene. Eco RI adapters were followed by  
digest with Xho I/Eco RI and ligated to pBluescript."

BASE COUNT 243 a 213 c 227 g 238 t 14 others  
ORIGIN

Query Match 96.6%; Score 766.8; DB 11; Length 935;  
Best Local Similarity 98.5%; Pred. No. 1.9e-209;

Matches 776; Conservative 7; Mismatches 4; Indels 1; Gaps 1;

```

OY 1 ccgcgaacttaccgctccacagctcattatattgttcgggaagctagagaat 60
DB 813 CCGGACMCTTATCCGCCCTCCACCGCTCTATTATTGTCGGGAGAGTAAAGT 754
OY 61 agtcgcagtaataatagtttcgcaacgttgtccatgtctaaagcaatcgtgtca 120
DB 753 AGTCCGCGCATTAATAGTTGGCCACAGTTGTCATTGCTACAGCATCGTGATCA 694
OY 121 cgcctcgttggtagtgatgctcattacacgtccgcgaagcgaagcttaca 180
DB 693 CGCTCGTCTTTGGTAGGCTTCATTCACCTCGGTTCCCAACGATCARGCGAGTTACA 634
OY 181 tgaatcccaatgtgtgcaaaaagcggttagctcctcgtccctccgaatcgttca 240
DB 633 TGATCCCATGCTGTGCAAAAAGCGGTACTCTCGTCCGCTCCGATCTGTGTCACA 574
OY 241 agtaagtt-gggcgaggtttatcaactcagttatgagcaacgtcgaatctcttac 299
DB 573 AGTAAGTTGGCGCGCAGTTATCMCTATGTTATGAGCAGCAGCAATTAATCTCTAC 514
OY 300 tgcatagcatacgttaagaatgctttctgtgactgtagtaactaaacaagtcattc 359
DB 513 TGTGATGCCATCCGTAAGATGCTTTTCTGTGACTGTGAGTACTAACCAAGTCAATCTG 454
OY 360 agaatagtatgtagcgccgagctgtcctcgtccgcgctcaatacaggaataacgc 419
DB 453 AGAATAAGTGTATGCGCGCAGCTGCTTCCCGCGGCTCAATACGAGTAATACCGC 394
OY 420 gccacatagcagaacttaaaagtctcatatttgaaaaagttctcggggcgaaact 479
DB 393 GCCACATAGCAAGACTTTAAAGTGTCTCATTTGAAACGTTCTTGGGCGCAAAACT 334
OY 480 ctcaagaatcttaacgctgttgaatccagttcgaatgaacacacgttgaccgaactg 539
DB 333 CTCAGAGATCTTACCGGCTGTGAGATCCAGTTCGATGTMCCCACTCGTGCACCACTG 274
OY 540 atttccagcatcttacttccacagcgtttctgggtgagcaaaaacaggaagcaaaa 599
DB 273 ATCTTCACGACATCTTTTACTTTCACCGGCTTCTGGGTGAGCAAAAACAGGAAGCAAAA 214
OY 600 tgcgcgaaaaaaggaagaagcgacacggaagaatgttgaatactactcctctt 659
DB 213 TGGCGCAAAAAGGAAGAATAGGGGCGACGCAAAATGTTGAATACTCTCTCTTTT 154
OY 660 tcaatattatgaagcattatcaggttattgtctcatgagcgaatataatttgaatg 719
DB 153 TCAATATTATTGAAGCATTTATTCAGGGTTATTGTCATGAGCGGATACATATTGTAATG 94
OY 720 tatttagaaaataaacaagaaggttcggcgacatctcccgaaaagtgcacactga 779
DB 93 TATTTAGAAAATATAACAATAATAGGGGTTCCGCGACATTTCCCGAAAAGTGCCACTAA 34
OY 780 cgtctaaag 787
DB 33 ATTGTAAAG 26

```

RESULT 2  
AL042026 841 bp mRNA EST 29-FEB-2000  
LOCUS DKEZp434E111\_r1 434 (synonym: htes3) Homo sapiens cDNA clone  
DEFINITION DKEZp434E111 5', mRNA sequence.  
ACCESSION AL042026  
VERSION AL042026.1 GI:5421372  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 841)  
AUTHORS Poustka, A., Klein, M., Mewes, H.W., Gassenhuber, J. and Wiemann, S.  
TITLE EST (Poustka, et al.)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Poustka A.J.  
Department Lehrach  
Max-Planck-Institute for Molecular Genetics  
Innestrasse 73, 14195 Berlin, Germany  
Tel: +49-30-84131623  
Fax: +49-30-84131128  
Email: poustka@mpg-berlin-dahlem.mpg.de  
This is the 5' sequence of the clone insert  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
sequenced by DKFZ (German Cancer Research Center,  
Heidelberg/Germany) within the cDNA sequencing consortium of the  
German Genome Project.  
No 5' sequence available.  
This clone (DKEZp434E111) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubneweg 6, 14059  
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.  
location/Qualifiers  
1. 841  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="DKEZp434E111"  
/clone\_1ib="434 (synonym: htes3)"  
/tissue\_type="testis"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/note="Vector: pSport1; Site\_1: NotI; Site\_2: SalI"

## FEATURES

## SOURCE

BASE COUNT 226 a 192 c 221 t  
ORIGIN

Query Match 74.0%; Score 741; DB 10; Length 841;  
Best Local Similarity 100.0%; Pred. No. 4.8e-202;  
Matches 741; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 ccgcgaacttaccgctccacagctcattatattgttcgggaagctagagaat 60
DB 741 CCGGACMCTTATCCGCCCTCCACCGCTCTATTATTGTCGGGAGAGTAAAGT 682
OY 61 agtcgcagtaataatagtttcgcaacgttgtccatgtctaaagcaatcgtgttca 120
DB 681 AGTCCGCGCATTAATAGTTGGCGCAAGCTGTGTCATTGCAAGGCAATCGTGATCA 622
OY 121 cgcctcgttggtagtgatgctcattacacgtccgcgaagcgaagcttaca 180
DB 621 CGCTCGTCTTTGGTAGGCTTCATTCACCTCGGTTCCCAACGATCARGCGAGTTACA 562
OY 181 tgaatcccaatgtgtgcaaaaagcggttagctcctcgtccctccgaatcgttca 240
DB 561 TGATCCCATGTTATGCAAAAAGCGGTACTCTCGTCCGCTCCGATCTGTGTCACA 502
OY 241 agtaagttggcgaggtttatcaactcagttatgagcaacgtcgaatctcttac 300
DB 501 AGTAAGTTGGCGCGCAGTTATCAGCTCATGATGATGAGCGACTGCAATTAATCTTACT 442
OY 301 gtcatagcatacgttaagaatgctttctgtgactgtagtaactaaacaagtcattcga 360

```

Db	441	GTATGCGCATCCGTAAGATGCTTTTCGTGTACTGTGAGTACTCAACCAAGTCATTTCTGA	382
OY	361	gaatagtgatgacgagcagcaggtgtgctctgtcccgcgctcaataagagataatccg	420
Db	381	GAATAGTGTATGGGGGACAGTGTGCTCTTGCCCGGCTCATATCGGATATATACCGG	322
OY	421	cccatagacgaacttaaaagtgtcactcaatggaacagttcttcggcgcgaaatc	480
Db	321	CCCATATGCGAAGACTTTTAAAGTGTCTCATTTGGAAAAGCGTTCTTGCGGCGCAAAATC	262
OY	481	tcaagatcttaacgctgttgagatcagtcogatgtaaccactcgtgaccccaactga	540
Db	261	TCAAGATTTTACCGCTGTTGAGATCCAGTTGATGCAATCCCACTCTGTGACCCAACTGA	202
OY	541	tcttcagcatcttctacttccagcagcgtttctgtgtgagcaaaaacaggaagcaaat	600
Db	201	TCTTCACCATCTTTTACTTTCACCAACGCTTCTGGGTAGCAAAAACGGAAGGCAAAAT	142
OY	601	gcgcgaaaaaaggaataagcgcgacacggaatgttgatatactatactctctctttt	660
Db	141	GCCGCAAAAAGSGAATTAAGGGGACACGGAATGTTGAATATCATATCTTCTCTTTT	82
OY	661	caatatattgaagcatttatacaggttatgttccatagaacggaatcatattgaaat	720
Db	81	CAATATATTGGAAGCATTTATTCAGGGTTATTTGTCATGACGGATGCATATTTGAATGT	22
OY	721	atttagaaaaataaacaata 741	
Db	21	ATTTAGAAAAATTAACAATA 1	

```

FEATURES             SOURCE
1.754
LOCUS                BG920379
DEFINITION           BG920379 754 bp mRNA
ACCESSION            BG920379
VERSION              BG920379.1 GI:14300855
KEYWORDS              EST.
SOURCE               house mouse.
ORGANISM             Mus musculus
AUTHORS              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
JOURNAL              Mammalia; Euteria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
COMMENT              NIH-MGC http://mgc.nci.nih.gov/.
                    Unpublished (1999)
                    Contact: Robert Strausberg, Ph.D.
                    Email: cgraphs-remail.nih.gov
                    Tissue Procurement: Jeffrey Green M.D.
                    CDNA Library Preparation: Life Technologies, Inc.
                    CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
                    DNA Sequencing by: Incyte Genomics, Inc.
                    Clone distribution: MGC clone distribution information can be
                    found through the I.M.A.G.E. Consortium/LNLN at:
                    http://image.llnl.gov
                    Plate: L1AM10916 rgw: o column: 08
                    High quality sequence start: 15
                    High quality sequence stop: 717.
                    Location/Qualifiers
                    1..754

```

```

/organism="Mus_musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone IMAGE:4954975
/clone_1b="NCI_CGAP_1aam6"
/sex="female_virgin"
/tissue_type="infiltrating ductal carcinoma"
/development="months"
/lab_host="H1D08"
/notes="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT
library constructed by Life Technologies. Investigator

```

providing samples: Jeffrey Green, M.D., NIH

BASE COUNT	201 a	174 c	166 g	213 t
ORIGIN				

Query Match	67.38;	Score 673.2;	DB 11;	Length 754;
Best Local Similarity	96.58;	Pred. No. 1.5e-182;		
Matches 710; Conservative	0;	Mismatches 23;	Indels 3;	Gaps 2

Oy	77	gttcgcgaacgcttgctgacattgctacaggaacgtgtgtgtcgaagctctgtctgttgcga	120
Db	11	CTTTGCCGAAGTGTGTGGCATTTGTAACAAGGACGTGTGTGCACGCTGTGCTTGGTA	70
Oy	137	tgactcattcagctccggttccccaaagatcaaggagatatacatgatcccccatggtgt	196
Db	71	TGGCTCATTTACAGTCCGGTGTCCCAACGATCAAGGGAGTTCATGATATCCCATGTTGT	130
Oy	197	gcataaagcggttagctcttcctcgttccctccgaatgtgttcgaagtaagtgtgcgcag	256
Db	131	GCATAAAGCGGTTAGCTCCCTCGGCTCTCCGATCGTTGTGCAAGTAAAGTGGCGCAG	190
Oy	257	tgatcaacatggtatagcagcagcactgtgataattctcttactgtcatgcatccgtaa	316
Db	191	TGTTATTCATGTAGTTATGGCAGACACTGGATATTCTCTTACTGTCAATGCATCCGTA	250
Oy	317	gatgctttctgtgactggtgtgagttaactcaaccgaatctcttgaagaatgttlatgggc	376
Db	251	GATGCTTTCTGTGTACTGTGACTGTGACTACCAACCAAGTCAATTCAGATAGTGATGTGGC	310
Oy	377	gacgcagttgctcttgcgcgcgcgcftcaatacggagataaacgcgcacatagcagaact	436
Db	311	GACCGAGTTGCTCTTGGCCCGCGCTCAATACGGGATATATACGGCGCCACATAGAGAACTT	370
Oy	437	taaaagtgcatacatltggaanaacgttcttcgtgggcgaaaactccaagatcttaccgc	496
Db	371	TAAAGTGCTACTCATTTGGAAGAGCTTCTTGCGGGCGCAAACTCTAGAGATTTACCGC	430
Oy	497	tgttgagatcagttcgaatgtaaccgaactcgttgacccaactgatcttcaagatcttca	556
Db	431	TGTTGAGATCCAGTTTCGATGTGAACCCACTCTGTGACCCAACTGATCTTTCAGCATCTTTTA	490
Oy	557	ctttacacgcgtttcttggtgtgagcaaaaacgaagaagcgaatgtccga - 'aaaaagg	614
Db	491	CTTTCCACCGCGTTCTGTGGGTGAGCAAAAACGGAAGGCAAAATGCCGCAACAAAGGGA	550
Oy	615	aataaaggcagcaggaagaatgttgaataactatactcttcttllcaatataatgaag	674
Db	551	ATPACGGCGACACAGGAATGTGAATCTATCTCTCTTTCAATATATTGAG	610
Oy	675	catltaacaggtatctgtctcaatgaagcggatataattgaatgtatattgaaaaaata	734
Db	611	CATTTATACAGGTTATTGTCTCATGAGCGGATACATATTATTAAGATATTTAAAGAAATTA	670
Oy	735	acaaatagggttccg - cgcacattccccggaagaatgtcacctcgaactcgaagaacca	793
Db	671	ACAAATAGGGGTTCCGCGCACATTTTCCCGAAAGTCCACCTAATTTGACAGCGTTA	730
Oy	794	ttatlatcatgacat 809	
Db	731	ATATTTGTTAAATTT 746	

RESULT	4
AL040542/c	
LOCUS	AL040542 767 bp mRNA
DEFINITION	DKFPP43411114_r1 424 (synonym: htess3) Homo sapiens cDNA clone
ACCESSION	DKFPP43411114 5', mRNA sequence.
VERSION	AL040542
KEYWORDS	AL040542.1 GI:5409490
SOURCE	EST.
ORGANISM	human.
	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 767)  
 AUTHORS Koehrer, K., Beyer, A., Mewes, H.W., Gassenhuber, J. and Wiemann, S.  
 TITLE EST (Koehrer, et al.)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Koehrer K  
 MIPS

Am Klopferplatz 18a D-82152 Martinsried, Germany

This is the 5' sequence of the clone insert  
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
 sequenced by BMFZ (Biomedical Research Center at the Charité,  
 Berlin/Germany) within the cDNA sequencing consortium of the German  
 Genome Project.

No sl sequence available.

This clone (DKFZp341i114) is available at the RZPD in Berlin.  
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

## FEATURES

source

1..767  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="DKFZp341i114"  
 /clone\_lib="434 (synonym: htes3)"  
 /tissue\_type="testis"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /note="Vector: pSPORT1; Site\_1: NotI; Site\_2: SalI"  
 BASE COUNT 219 a 158 c 168 g 221 t 1 others  
 ORIGIN

Query Match 65.0% Score 650.6; DB 10; Length 767;  
 Best Local Similarity 99.4% Pred. No. 4.6e-176;  
 Matches 653; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

131 ttgtatggtctcatcgcctccggttcccaagatcaagcgagttatcatcccca 190  
 765 ttgcatggtcttcatcagctccggttcccaagatcaagcgagttatcatcccca 706  
 191 ttttgcacaaacgggttgcctcctcgtcccgatcgtgtcagaagtgttg 250  
 705 tctttgcaaaaacgggttgcctcctcgtcccgatcgtgtcagaagtgttg 646  
 251 ccgagtggttactacatcgttgcagcagtcatactcttactgcatgcat 310  
 645 ccgagtggttactacatcgttgcagcagtcatactcttactgcatgcat 586  
 311 ccgtaagatgctttctgtgactggtgactcaaccaagtcattcgaatagtta 370  
 585 cctgtagatgctttctgtgactggtgactcaaccaagtcattcgaatagtta 526  
 371 tgcgagcagcagtgctcttcccgccggtcaatacggagtaaacggcgacataga 430  
 525 tgcgagcagcagtgctcttcccgccggtcaatacggagtaaacggcgacataga 466  
 431 gaacttaaaagtgcatactatggaaaacgcttctcggggcgaaaactccaagatct 490  
 465 gaacttaaaagtgcatactatggaaaacgcttctcggggcgaaaactccaagatct 406  
 491 tacgccttgaagatcgaatgtaaacacactcgtgacccaactgactcgaat 550  
 405 tacgccttgaagatcgaatgtaaacacactcgtgacccaactgactcgaat 510  
 551 ctttacttaccagcgcttctcgggtgagcaaaaacgaaagcaaaaactccgcaaaaa 610  
 345 ctttacttaccagcgcttctcgggtgagcaaaaacgaaagcaaaaactccgcaaaaa 286  
 611 agggaaataagggcgacacggaataatgtgaatactactactctcctttcaataat 670  
 285 agggaaataagggcgacacggaataatgtgaatactactactctcctttcaataat 226  
 671 gaagcattatcaggttatgtcctatgagcgagatacatatttgaaatattagaaaa 730

|||||  
 Db 225 GAACCATTTATCAGGGTATTGCTCATGACCGCATACATTTTGAATATTGAAAA 166  
 Qy 731 ataaacaataaggggttcgcgcacattcccccgaagaatgcacactgacgtcga 787  
 Db 165 ATAACAATAAGGGGTTCGCCGACATTTCCCGAATAAGTCCACCTAATATTGAG 109

## RESULT

5

AL598527 693 bp mRNA

LOCUS DKFZp31301920.r1 313 (synonym: h1cc2) Homo sapiens cDNA clone

DEFINITION DKFZp31301920 5', mRNA sequence.

ACCESSION AL598527

VERSION AL598527.1 GI:15161218

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

EST (Duesterhoeft, et al.)  
 Unpublished (1999)  
 Contact: Duesterhoeft A  
 MIPS

Am Klopferplatz 18a D-82152 Martinsried, Germany

This is the 5' sequence of the clone insert  
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
 sequenced by Qigen (Hilden/Germany) within the cDNA sequencing  
 consortium of the German Genome Project.

No sl sequence available.

This clone (DKFZp31301920) is available at the RZPD in Berlin.  
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

## FEATURES

source

1..693  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="DKFZp31301920"  
 /clone\_lib="313 (synonym: h1cc2)"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /note="Vector: pTriplex2; Site\_1: SfiI; Site\_2: SfiIb;  
 cDNA-collection"  
 BASE COUNT 173 a 167 c 156 g 194 t 3 others  
 ORIGIN

Query Match 63.7% Score 638; DB 10; Length 693;  
 Best Local Similarity 99.7% Pred. No. 1.9e-172;  
 Matches 638; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

362 aatggtatgacggcgacgagtgctcgtccggcggtcaataagggataacgcgc 421  
 Db 685 AATGATGATGCGGCGACGAGTGTGCTTGGCCGCGCTCAATACGTAATATACCGCCG 626  
 422 cacaatacagaacttaaaagtgtcactcaatcgtgaaacgcttctcggggcgaaaactct 481  
 Db 625 CACATAGAGAGANTTTAAAGTGTCTATCTATGGAAGAGTCTTCTGGGGCAAAACTCT 566  
 482 caagatcttaacgcgtgtgagatcgaatgtaaacacactcgtgacccaactgat 541  
 Db 565 CAAGATCTTACCGCGTGTGATTCAGATGCAATGCAACCACTGTCGACCACTGAT 506  
 542 cttcagcatttacttaccagcgcttctcgggtgagcaaaaacggaaggaagaaatg 601  
 Db 505 CTTGAGATCTTTACTTTTACCGAGGTTTCTGGGTGACCAAAACAGCAAAATG 446  
 602 ccgcaaaaagggaaataagggcgacacggaataatgtgaatactactactccttttc 661

Db 445 CCGCAAAAGGGAATAGGCGACAGGAATGTTGATCTACTCTCTCTTTTC 386  
 Qy 662 aatattatgaagcattatcagaggtatctcctcagagcgatataatattatga 721  
 Db 385 AATATTATTGAAGCATTTATTCAGGGTTATTCATGAGGGGATACATTTTGAATGTA 326  
 Qy 722 tttaaaaaataaacaantggggttcgcgcacattcccccgaagaagtgcacactgacg 781  
 Db 325 TTTTAAAGAAATTAACAAATAGGGGTTCCGGCCACATTTCCCGGAAGGTGCACACTGACG 266  
 Qy 782 tctaaagaacattatcattcagatatacctataaataagcgatatacagagcgct 841  
 Db 265 TCTAAGAAACATTTATTCATGACATTAACCTATAAAAATAGCGGTATACGAGGCGCT 206  
 Qy 842 ttcgctcgcgcgttcggtgtagcagtgaaacattcagacatgcagctcccgaga 901  
 Db 205 TTTGCTCTCCGGGTTGGTGTATGAGCGGTGAACCTCTGACACATGACGCTCCGGGAGA 146  
 Qy 902 cgcgtcacactgctcgttaagcagatgcgcgagagacaagcccgatcagagcgctcag 961  
 Db 145 CGGTACACACTTGTCTGTATAGCGGATGCCGGAGACAGAACCCGCTCAGGCGCTCAG 86  
 Qy 962 cgggtgtcggcggtgtcggggtcgtgcttaactatgcgcgc 1001  
 Db 85 CGGGTGTGGCGGGTGTGGGCGCTGACTTAACTATGCGGC 46

RESULT 6  
 AL042640/c 688 bp mRNA EST 29-FEB-2000  
 LOCUS AL042640.1 GI:5422090  
 DEFINITION DKFZP434M0421.1 434 (synonym: htes3) Homo sapiens cDNA clone  
 AL042640  
 VERSION AL042640.1 GI:5422090  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens.  
 REFERENCE Blum, H., Bauersachs, S., Mewes, H.W., Gassenhuber, J. and Wiemann, S.  
 1 (bases 1 to 688)  
 TITLE EST (Blum, et al.)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Blum H  
 MIPS Am Klopferstr. 18a D-82152 Martinsried, Germany  
 This is the 5' sequence of the clone insert  
 This is the 5' sequence of the clone insert  
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
 Research Center (DKFZ), Email: s.wiemann@dkfz-heidelberg.de;  
 sequenced by LMU (Ludwig Maximilians University,  
 Munich/Germany) within the cDNA sequencing consortium of the German  
 Genome Project.  
 No si sequence available.  
 This clone (DKFZP434M0421) is available at the RZPD in Berlin.  
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES  
 source  
 1..688  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="DKFZP434M0421"  
 /clone\_id="434 (synonym: htes3)"  
 /tissue\_type="testis"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /note="Vector: pSPORT1; Site\_1: NotI; Site\_2: SalI"

BASE COUNT 190 a 150 c 158 g 190 t  
 ORIGIN  
 Query Match 63.6%; Score 636.6; DB 10; Length 688;  
 Best Local Similarity 99.4%; Pred. No. 4.7e-172;  
 Matches 639; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 145 ttcagctcgcgttcccaagatcagagcgatatacagatcccccagtgctgcaaaaaa 204  
 Db 688 TTTACCTCCGGTTCCTCCCAAGATCAAGGAGATTACATGATCCCCCATGTTGTCACAAAAA 629  
 Qy 205 ggcgttagctcttcgcgtcccgatcgtgttcgagaagtaagtggcgcagtgatca 264  
 Db 628 GCGGTACTCTCTTGGTCCGTCGAGTGGTTCAGAGTAAGTTGGCGGAGTGTATCA 569  
 Qy 265 ctcatggttaaggacagcagcagcagcagcagcagcagcagcagcagcagcagcagc 324  
 Db 568 CTCATGTTATGCGACGACTGCTCATTAATTCCTTACTGTATGATGCCATCCGTAAGATGCTTT 509  
 Qy 325 tctgtgacgtgtgagttactcaaccaagcagcagcagcagcagcagcagcagcagcagc 384  
 Db 508 TCTGTACTGTGAGTACTCAACCAAGCATTCATTCAGATAGTATGTCGCGCACGAGT 449  
 Qy 385 tgcctcgcgcgcgtcactaagcagcagcagcagcagcagcagcagcagcagcagcagc 444  
 Db 448 TGTCTTCCCGCGGCAATATACGGATTAATACCGGCGCACATAGCAAGATTTAAAGTG 389  
 Qy 445 ctcatcattggaagaacgttctcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 504  
 Db 388 CTCATATTGGAAGAAAGCTTCTTCGGGCGCAAAACTCTCAGATCTTACCGCTGTGAGA 329  
 Qy 505 tccagctcagtgtaaccacatcgtgcacccaactgactcagcagcagcagcagcagcagc 564  
 Db 328 TCCAGTTGCATGTAAACCCACTGTCACCACTGATCTTCAGCATCTTTACTTTCAC 269  
 Qy 565 agcgttctcgtgtgagcaaaaacagcagcagcagcagcagcagcagcagcagcagcagc 624  
 Db 268 AGGTTTCTGGGTGACCAAAACAGGAGGCAAAATGCGGCAAAAAAGGAATAGGCGC 209  
 Qy 625 acaaggaatgtgtaactcactcctccttcttcaatattatgaagcattatcaag 684  
 Db 208 ACAGGAATGTTGAATACATCATCTTCTCTTCAATATTATGAGCATTTATGAG 149  
 Qy 685 gattatgtcctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 744  
 Db 148 GGTATTGTCTCAAGACGAGATACATATTGAAATGATTTAGAAAAATTAACCAATAGG 89  
 Qy 745 gtccgcgcacattcccgaaagtgcacactgacgctcag 787  
 Db 88 GTTCCGCGCACATTTCCCGCAAAAGTCCACCTAAATTTGAAAG 46

RESULT 7  
 BE569261 914 bp mRNA EST 15-AUG-2000  
 LOCUS BE569261.1 NCL\_CGAP\_Mam6 Mus musculus cDNA clone IMAGE:370859 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BE569261  
 VERSION BE569261.1 GI:9812981  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus.  
 REFERENCE Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 914)  
 TITLE NIH-MGC http://mgi.nci.nih.gov/  
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cga@bhs.fda.gov  
 Tissue Procurement: Jeffrey Green M.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LINL at:  
 http://image.llnl.gov  
 plate: LHM8928 row: m column: 20  
 High quality sequence stop: 685.

## FEATURES

source  
 Location/Qualifiers  
 1. 914  
 /organism="Mus musculus"  
 /strain="FVB/N"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:3708859"  
 /sex="female, virgin"  
 /tissue\_type="infiltrating ductal carcinoma"  
 /dev\_stage="5 months"  
 /lab\_host="DH10B"  
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site-1: SalI; Site-2: NotI; Cloned unidirectionally. Primer: Oligo dt. library constructed by Life Technologies. Investigator providing samples: Jeffrey Green, M.D., NIH"  
 BASE COUNT 275 a 207 c 199 g 233 t  
 ORIGIN

Query Match 60.7%; Score 607.2; DB 10; Length 914;  
 Best Local Similarity 98.6%; Pred. No. 1,4e-163;  
 Matches 623; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

157 tcccaacgaatcaagcagatcatcagatcccccagtggtgcaaaaagcggttagctcc 216  
 1 tccgaacgatcaagcagatcatcagatcccccagtggtgcaaaaagcggttagctcc 60  
 217 ttcggtcccgatcggtgcaagaagtaagtcgcccagtggtatcatcatggtatg 276  
 61 ttgcggtcccgatcggtgcaagaagtaagtcgcccagtggtatcatcatggtatg 120  
 277 gaagcagatgataatcct 336  
 121 gcagcagatgataatcct 180  
 337 gactctcaacaaagtaattctgagaagtagtatacgcgagcagatgctctgccc 396  
 181 gagtctcaacaaagtaattctgagaagtagtatacgcgagcagatgctctgccc 240  
 397 ggtcaaaacaggaagtaacacgagcagcagcagcagcagcagcagcagcagcagc 456  
 241 gcgtcaaaacaggaagtaacacgagcagcagcagcagcagcagcagcagcagcagc 300  
 457 aacggtctctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 516  
 301 aacggtctctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 360  
 517 taaccacagtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 576  
 361 taaccacagtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 420  
 577 tgaagcaaaacaggaagtaacacgagcagcagcagcagcagcagcagcagcagcagc 636  
 421 tgaagcaaaacaggaagtaacacgagcagcagcagcagcagcagcagcagcagcagc 480  
 637 tgaagcaaaacaggaagtaacacgagcagcagcagcagcagcagcagcagcagcagc 696  
 481 tgaagcaaaacaggaagtaacacgagcagcagcagcagcagcagcagcagcagcagc 540  
 697 atgagcagatatacttgaatgattatgaaataaataaataaataaataaataaataa 756  
 541 atgagcagatatacttgaatgattatgaaataaataaataaataaataaataaataa 600  
 757 -ttccccgaaaagtgcaactgacgtcgaag 787  
 601 tttttcccgaaaagtgcaactgacgtcgaag 632

RESULT 8  
 AL042540 702 bp mRNA EST 29-FEB-2000  
 LOCUS AL042540.1 434 (synonym: htes3) Homo sapiens cDNA clone  
 DEFINITION DKFp434I1621 5', mRNA sequence.

ACCESSION AL042540  
 VERSION AL042540.1 GI:5421995  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 702)  
 AUTHORS Blum, H., Bauersachs, S., Mewes, H.W., Gassenhuber, J. and Wiemann, S.  
 TITLE EST (Blum, et al.)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Blum H  
 MISC

Am Klopferspitz 18a D-62152 Martinsried, Germany  
 This is the 5' sequence of the clone insert  
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
 sequenced by LMU (Ludwig Maximilians University,  
 Munich/Germany) within the cDNA sequencing consortium of the German  
 Genome Project.  
 No sl sequence available.  
 This clone (DKFp434I1621) is available at the RZPD in Berlin.  
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
 Berlin-Charlottenburg, GERMANY; Email: clone@rzd.de.

## FEATURES

source  
 Location/Qualifiers  
 1. 702  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="DKFp434I1621"  
 /clone\_id="434 (synonym: htes3)"  
 /tissue\_type="testis"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /note="Vector: pSPORT1; Site-1: NotI; Site-2: SalI"  
 BASE COUNT 185 a 159 c 164 g 192 t 2 others  
 ORIGIN

Query Match 59.1%; Score 591.4; DB 10; Length 702;  
 Best Local Similarity 96.2%; Pred. No. 4.7e-159;  
 Matches 637; Conservative 0; Mismatches 22; Indels 3; Gaps 3;

129 gtttgatgagctcattcagctccggttcccaagcagcagcagcagcagcagcagcagc 187  
 701 gtttgatgagctcattcagctccggttcccaagcagcagcagcagcagcagcagcagc 642  
 188 ccaatggtgcaaaaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 246  
 641 ccaatggtgcaaaaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 582  
 247 ttggcagcagtgatcattcagcagcagcagcagcagcagcagcagcagcagcagcagc 305  
 581 ttggcagcagtgatcattcagcagcagcagcagcagcagcagcagcagcagcagcagc 522  
 306 gccatcgtaagatgcttctctgtagcagcagcagcagcagcagcagcagcagcagcagc 365  
 521 gccatcgtaagatgcttctctgtagcagcagcagcagcagcagcagcagcagcagcagc 462  
 366 gttgatgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 425  
 461 gttgatgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 402  
 426 tagcagaacttaaaagtgctcattcagcagcagcagcagcagcagcagcagcagcagc 485  
 401 tagcagaacttaaaagtgctcattcagcagcagcagcagcagcagcagcagcagcagc 342  
 486 gacttaacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 545  
 341 gacttaacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 282  
 546 agcattcttattcaccagcagcagcagcagcagcagcagcagcagcagcagcagcagc 605  
 281 agcattcttattcaccagcagcagcagcagcagcagcagcagcagcagcagcagcagc 222





```

Db 598 TCTTCCCGGCTGAATACGGATAATACCGCCACATAGCAGAACTTT-AAAGTCTC 540
Qy 448 atcatgtgaaacgtctctcgggacgaacatcgaagatcttaccgc--tgtagagt 505
Db 539 ATCATTTGGAAAAAGCTTCTCGGGCGGAAACCTCTCAAGATCTTACCAGCTGTTGAGAT 480
Qy 506 ccaatttgatgtaaccacacgtcgtagcacaacatgatcttcgaatcttacttcaacca 565
Db 479 CCAATTTGATGTAACCCACATCGTGCACCCAACTGATCTTCAGATCTTACTTTCACCA 420
Qy 566 gcgttctgggtgagcaaaacaggaagaaatccgcaaaaaaaggaaataaggcgca 625
Db 419 GCGTTTCTGGGTAGCAAAAAGGAAAGGCAAAATGCCCCAAAAGGGAATAGGCGCA 360
Qy 626 caaggaatgtgtaactcaatctctccttcttcaatataatgaacatlaacagg 685
Db 359 CACGGAATGTGTAATACATACATCTCTCTCTTTTCAATATTATTAGAACATTTATCAG 300
Qy 686 gttatgtctcatgagcgaatcatatttgaaatgtatttaagaataaacaataaggag 745
Db 299 GTTATTGTCTCATGACGGGATACATTTGAAATGATTGAAATAAACAATAAGAGGG 240
Qy 746 ttccgcgcacatttcccgaaagatgcacacgtcgcgtctaaagaacatlaatcaga 805
Db 239 TTCGCGGACATTTCCCGAAAAGTCCACCTGACGTCTAAGAACATTAATATCATGA 180
Qy 806 catlaaccataaaatagcgatcaacgaagcccttcgtctcgcggttctgtagatg 865
Db 179 CATTAACTCATTAATAATAGCGCTATACAGAGCCCTTCTGCTCGCGGTTTGGTGTAG 120
Qy 866 acggtgaaacactctgacacatgcagctcccggaagaggtacagcttctgtagaagcg 925
Db 119 ACGGTGAAACCTCTGACACATGCGAGCTCCCGAGACGNTACAGCTTGTCTGTAACGG 60
Qy 926 atgcgggagcagacagccgcgtcagggcgcg 957
Db 59 ATGCGGGAGCAGNACGCCGTCACATCCCG 28

```

```

RESULT 11
AC000364/c 793 bp DNA GSS 06-FEB-1999
LOCUS Homo sapiens genomic DNA, 21q region, clone: PQ119B8X30, genomic
DEFINITION survey sequence.
ACCESSION AG000364
VERSION AG000364.1 GI:2579172
KEYWORDS GSS.
SOURCE Homo sapiens DNA, clone: PQ119B8X30.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 793)
AUTHORS Hattori, M., Ishii, K., Shiba, T. and Sakaki, Y.
TITLE Direct Submission
JOURNAL Submitted (31-OCT-1997) to the DDBJ/EMBL/GenBank databases.
Masahira Hattori, Kitasato University, Department of Science, JST
Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan
(E-mail: hattori@hgc.ims.u-tokyo.ac.jp, Tel: 0427-78-9732,
Fax: 0427-78-9561)
2 (bases 1 to 793)
REFERENCE Hattori, M., Ishii, K., Shiba, T. and Sakaki, Y.
TITLE Homo sapiens genomic DNA, chromosome 21q
JOURNAL Published Only in Database (1997) in press
FEATURES
Location/Qualifiers
1..793
/organism="Homo sapiens"
/db xref="taxon:9606"
/chromosome="21"
/clone="PQ119B8X30"
/map="21q"

```

```

BASE COUNT 205 a 185 c 207 g 180 t 16 others
ORIGIN

```

```

Query Match 55.9%; Score 559.2; DB 13; Length 793;
Best Local Similarity 97.7%; Pred. No. 8, 6e-150;
Matches 594; Conservative 0; Mismatches 11; Indels 3; Gaps 3;

```

```

Qy 1 ccttgcaacttaccgcgtccatccagctcattatattgttgcgggaagctagagtaag 60
Db 634 CCGCACTTATCCCGCTCCATCCAGTCTATTATTTGCGCGGAAGCTAGACTAGT 575
Qy 61 agtctgcagttatagatgttcgcaacgtctgtgcatgtgcaagg--catcgtgtgtc 119
Db 574 AGTTGGCAGTTATATGTTTGGCAACGTTTGGCATTGCTACAGGCCATCGTGTGTC 515
Qy 120 acgctgtgtgtgtgtagtggtcattcaagctccggttcccaagatcaagcgagttac 179
Db 514 ACCTCGCTGCTTGGTATGGCTTCATTCAGCTCCGCTGCCAAGCATCAAGCGAGTTAC 455
Qy 180 atgaccccaatgtgtgcaaaaaagcggttagctcccttggtctccgatcgtgtcag 239
Db 454 ATGATCCCCCATGTTTGGCAAAAAGCGGTGATGCTTTCGCTCCGATCGTTGTAC 395
Qy 240 aagtaagttggcgcagtgltatcaatcagtggtttagcaagcatgataatctctac 299
Db 394 AAGTAGTGGCGCGCAGTGTATCATCTCATGCTNATGGCAGCATGATTAATTTCTTAC 335
Qy 300 tttcatgcatccgtgaagatgtcttctgtgactgtgtgtagtactcaacaaatcatctg 359
Db 334 TGTCAATGCCATCGTAATAGTCTTTCGTGACGTGTAGTACCAACAAGATCATCTGTG 275
Qy 360 aagaatagltatgcgcgacagcagttgtctctgcgcgcggtcgaataggaataatccgc 419
Db 274 AGAATAGTGTATGGCGGACGACGAGTGTCTTCCGCCGCTCAATACGGATTAATACCG 215
Qy 420 gccacatgacgaactttaaagtgctcaatcattggaaa--cgttctcgggacgaagac 478
Db 214 GCCAATAGCAAACTTTAAAGTCTTTCATTTGAAANNCTGTTTTCGCGCGGAANN 155
Qy 479 t-clcaagatctaccgcgtgttgagatccagttcgaatgtrtaaccactgtgtgacccaa 537
Db 154 TNCNTAAGATCTTATCCGCTGNTGAGATCCAGTTGATGACCCACTGTCGACCCAAAC 95
Qy 538 tgatcttaagcatcttacttcaacagcgttctgtgtgagcaaaaaaggaagca 597
Db 94 TGATCTTCAGCATCTTATTCTTACACGAGTTTCNGGTCGNCAAAACAGGAAGCA 35
Qy 598 aatgcgcg 605
Db 34 AATGCTCT 27

```

```

RESULT 12
AC009765/c 696 bp DNA GSS 14-APR-1999
LOCUS Homo sapiens genomic DNA, 21q region, clone: PBG4SPN15, genomic
DEFINITION survey sequence.
ACCESSION AC009765
VERSION AC009765.1 GI:3289751
KEYWORDS GSS.
SOURCE Homo sapiens DNA, clone: PBG4SPN15.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 696)
AUTHORS Hattori, M., Ishii, K., Toyoda, A., Shiba, T. and Sakaki, Y.
TITLE Homo sapiens genomic DNA, chromosome 21q
JOURNAL Published Only in Database (1998) in press
REFERENCE 2 (bases 1 to 696)
AUTHORS Hattori, M., Ishii, K., Toyoda, A., Shiba, T. and Sakaki, Y.
TITLE Direct Submission
JOURNAL Submitted (06-JUL-1998) to the DDBJ/EMBL/GenBank databases.
Masahira Hattori, Kitasato University, Department of Science, JST
Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan

```



Qy	Db	Qy	Db
691	536	751	595
tgctccatgaagcggatgacatcttgcagtgtatcttagaagaataaacacaaataggggtctcg	tgtctcattgacgcgatgacatttttgatgtgttttagaanaataataacaaattaggggtt-cg	cgcacattctccgcgaaaaatgc	cgcacattttcccggaagaatgc
750	594	772	616

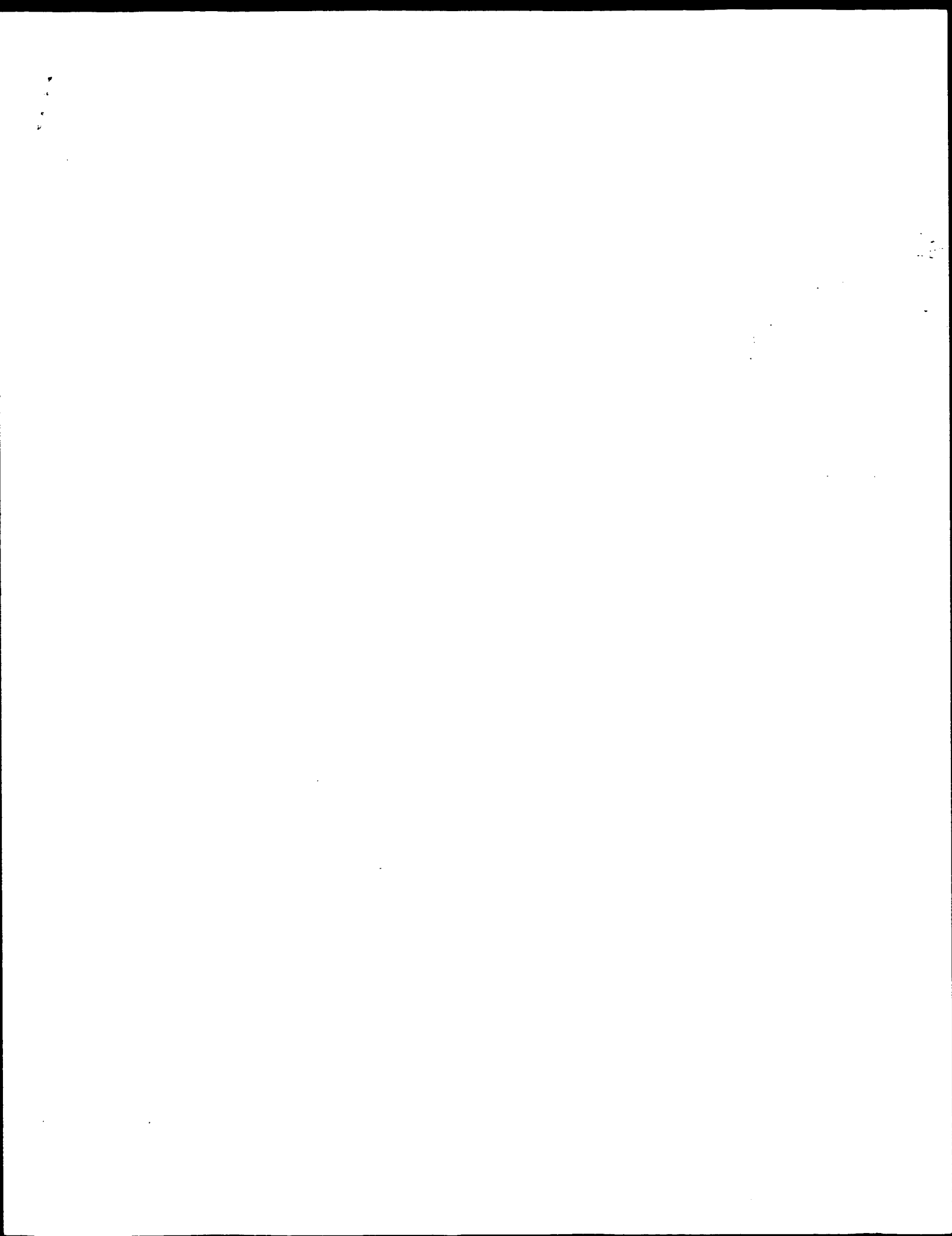
FEATURES	SOURCE
Location/Qualifiers	
1. .801	
/organism="Mns_musculus"	
/strain="FVB/N"	
/db_xref="taxon:10090"	
/clone_image="IMAGE:3709789"	
/clone_lib="NCI-CCAP_Mam6"	
/sex="female, Virgin"	
/tissue_type="infiltrating ductal carcinoma"	
/dev_stage="5 months"	
/lab_host="DH10B"	
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigator providing samples: Jeffrey Green, M.D., NIH"	
BASE COUNT	
213 a	194 c 177 g 217 t
ORIGIN	

Db	178	CCGATCGCTTGCAAGAAAGTGTGGCCGCACTGTATTACACATCAAGTGTATATGGCCAGCACTG	237
Qy	286	cataattctcttactgcatgcctccgtgaagaatgcttctctgactggtagaattaca	345
Db	238	CATATTCTCTTACTGTCATCCATCCGTAGATAGCTTTTCTGTGACTGTGTACTCA	297
Qy	346	accaagtcattctlgagaatagtagtatagtgacgacccagcttgctcttgccggcgataa	405
Db	298	ACCAAGTCATCTGTGAGATAGTGTATGTGGCGGACCCGAGTTGCTTTCGCCGGCTCAATA	357
Qy	406	cggagataataccggyccaataagaagaacttaa-aagtgctatactatggaaacgttc	464
Db	358	CGGGATATATACCGGCCCATATGAGAACTTTAAACAAGTGTATATGGAATAAGCTTC	417
Qy	465	ctcgaggcgaaaactctcaagaagatccttacgcgtgttgagatccagctgataccac	524
Db	418	TTTCGGGGCGAANAACCTCAAGATGCTTAAACGCTGTGAGATCCAGTTCGATGTAAACCAC	477
Qy	525	ctgfgacccaacatgaltcctcaagatccttacttcaaccaagcttctgtytgagaaa	584
Db	478	TCTGTCCACCCAAATGATCTTCAGATCTTWTATCTTTCACACGCTTTCTGTGGTAGGAAA	537
Qy	585	aa--caggaagcgaaaatctgcgcgcaaaaaagggaaataa--ggcgcgacaggaatatgtgaa-	640
Db	538	AAACGAGAAAGGCCAAATAATGCCGCAAAAAGGAGATTAACGGCGACAGGAATGTGTAAAC	597
Qy	641	-tactcaatacctctcccttttccaa--tatattgaagaactt-ataagggtatattgctc	696
Db	598	TACTCCATACGCTTTCCTTTTTCAAAGTATTATATGCAAGAGATTTCATACGGCTTATGGCTC	657
Qy	697	atgagc-ggataacatatttgaat-gtatatttgaanaaaataaacaaa--taagygltccgg	752
Db	658	ATGACCGGGATTCATATTGTGAAGCTACTGTAGAACAAATTAAACAACAATAAGGGGTACCGCG	717
Qy	753	caactttcccgaaaaagtgc 772	
Db	718	CACATATCCCGAAAAGTGC 737	

```
/clone_lib="434 (synonym: htes3)"  
/tissue_type="testis"  
/dev_stage="adult"  
/lab_host="DH10B"  
/note="vector: pSport1; Site_1: NotI; Site_2: SalI"  
BASE COUNT      152 a      123 c      124 g      155 t  
ORIGIN
```

```
Query Match      54.5%; Score 545.6; DB 10; Length 554;  
Best Local Similarity 99.3%; Pred No. 6.5e-146;  
Matches 548; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
  
QY 226 ccgacgtgtgtcagaagaagttggtccgagtgatcatcactatgttlatgcaagcactg 285  
    |||||||  
DB 554 ccgacgtgtgtcagaagaagttggtccgagtgatcatcactatgttlatgcaagcactg 495  
  
QY 286 cataattctctactgtcatgcacatccglaagaatgcttcttgactggtgaagtaactca 345  
    |||||||  
DB 494 cattaattctctactgtcatgcacatccglaagaagcttcttgactggtgaagtaactca 435  
  
QY 346 accaagtcattctgagaatagtgatgctggtccgagttgcttgcggtgcaata 405  
    |||||||  
DB 434 accaagtcattctgagaatagtgatgctggtccgagttgcttgcggtgcaata 375  
  
QY 406 cgggataataccgagccacatagcagaactttaaaagtgatcatcatggaacgttct 465  
    |||||||  
DB 374 cgggataataccgagccacatagcagaactttaaaagtgatcatcatggaacgttct 315  
  
QY 466 tcggggcggaanaactcaaggacttaccgctgttgagatccagttcgatgtaaccact 525  
    |||||||  
DB 314 tcggggcggaanaactcaaggacttaccgctgttgagatccagttcgatgtaaccact 255  
  
QY 526 cgtgaccccaactgacttctgagcatcttcttaccaggttctcgttgagcaaaa 585  
    |||||||  
DB 254 cgtgaccccaactgacttctgagcatcttcttaccaggttctcgttgagcaaaa 195  
  
QY 586 acaaggaagcaaaatgcgcgcdaaaaggaataaagcgacacgaaatgttgaatactc 645  
    |||||||  
DB 194 acaaggaagcaaaatgcgcgcdaaaaggaataaagcgacacgaaatgttgaatactc 135  
  
QY 646 ataactcttccttcaaatattatgaagcattatcaaggttattgtcctcatgagcgga 705  
    |||||||  
DB 134 ataactcttccttcaaatattatgaagcattatcaaggttattgtcctcatgagcgga 75  
  
QY 706 tacataattggaatgtaattagaataaacaabaataggggttcgcgcacatttcccgga 765  
    |||||||  
DB 74 tacataattggaatgtaattagaataaacaabaataggggttcgcgcacatttcccgga 15  
  
QY 766 aaagtgccacct 777  
    |||||||  
DB 14 aaagtgccacct 3
```

Search completed: January 8, 2002, 15:13:16  
Job time: 4792 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 8, 2002, 17:29:38 : Search time 428.83 Seconds  
(without alignments)  
1999.219 Million cell updates/sec

Title: US-09-740-211-14\_COPY\_1\_1000  
Perfect score: 1000  
Sequence: 1 cgcacctgacgagcagctgcg.....cagacctatggagcccta 1000

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 10%  
Listing first 45 summaries

Database :  
1: /SIDS2/gcgdata/geneseq/geneseq/NA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq/NA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq/NA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq/NA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq/NA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq/NA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq/NA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq/NA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq/NA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq/NA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq/NA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq/NA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq/NA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq/NA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq/NA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq/NA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq/NA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1000	100.0	4999	21	Recombinant adeno
2	1000	100.0	4999	22	Human factor VIII
3	736.2	73.6	11933	21	Recombinant adeno
4	736.2	73.6	11933	22	Human factor VIII
5	640	64.0	7944	22	Plasmid DL26 encod
6	593.2	59.3	12445	21	Vector HSOeneo fo
7	592.2	59.2	5035	18	Factor VIII-d8695-
8	590.6	59.1	4832	19	Human factor VIII
9	590.6	59.1	4832	19	Human factor VIII
10	590.6	59.1	8967	17	Factor-VIII full-1
11	590.6	59.1	8967	22	Human factor VIII

12	590.6	59.1	8975	6	AA50054	Human factor VIII
13	590.6	59.1	8975	21	AA38604	Human full-length
14	590.6	59.1	9029	22	AA60309	Human factor VIII
15	590.6	59.1	9068	19	AAV15359	Human factor VIII
16	590.6	59.1	9080	19	AAV19580	Human factor III e
17	589	58.9	9009	14	AA050185	Human factor VIII
18	589	58.9	9009	18	AA61548	Factor VIII:C (Arg
19	589	58.9	9009	19	AAV25810	Human factor VIII
20	589	58.9	9009	20	AAV91162	Human factor VIII
21	589	58.9	9009	22	AAV90508	Human factor VIII
22	587.2	58.7	9164	20	AAV82259	Beta-domain delete
23	587.2	58.7	11846	20	AAV82261	Factor VIII protei
24	587.2	58.7	12022	20	AAV82260	Factor-VIII CDNA-
25	586	58.6	6300	17	AAV03571	Factor VIII CDNA 1
26	586	58.6	8241	9	AAV81439	CDNA sequence enco
27	586	58.6	8241	9	AAV81096	CDNA sequence enco
28	585.8	58.6	9009	19	AAV18884	Human sapiens facto
29	585	58.5	7053	18	AAV1357	Sequence encoding
30	584.2	58.4	7440	7	AAV60689	Factor VIII:C codi
31	583.4	58.3	4830	9	AAV81544	Human factor VIII-
32	583.4	58.3	5094	21	AAV49231	DNA construct HSO/
33	583.4	58.3	7056	15	AAV66615	Sequence of human
34	581.8	58.2	4629	20	AAV88293	Human factor VIII
35	580.2	58.0	4629	16	AAV76016	B-domain deleted F
36	579.2	57.9	9354	18	AAV73164	CDNA encoding huma
37	566.6	56.7	7440	6	AAV50375	DNA sequence encod
38	565.8	56.6	4373	20	AAV82258	Beta-domain delete
39	528	52.8	4272	9	AAV80447	Modified factor VI
40	528	52.8	4275	10	AAV80446	Modified factor VI
41	528	52.8	4275	9	AAV80444	DNA encoding 740 A
42	528	52.8	4545	9	AAV73339	Modified factor VI
43	521.8	52.2	9009	19	AAV12112	Human factor-VIII
44	521.8	52.2	9009	19	AAV12112	Human factor-VIII
45	495.4	49.5	7914	22	AAV84648	Homo sapiens facto

## ALIGNMENTS

RESULT 1	AAV00122	standard; DNA; 4999 BP.
ID	AAV00122	
XX	AAV00122;	
AC	31-JUL-2000	(first entry)
XX		
DT		
DE	Recombinant adeno associated vector construct, pV4.JCF8deltaB.	
XX		
KW	Recombinant Adeno Associated Vector; PAAV; pV4.JCF8deltaB; promoter;	
KW	human factor VIII; hVfVIII; Efilalpa; human elongation factor-1alpha;	
KW	human growth hormone; hGH; ITR; inverted terminal repeat; haemophilila;	
KW	gene therapy; ds.	
XX		
OS	Adeno associated virus.	
XX		
PN	WO200023116-A1.	
XX		
PD	27-APR-2000.	
XX		
PF	19-OCT-1999;	99WO-US24495.
XX		
PR	20-OCT-1998;	98US-0104994.
PR	24-MAR-1999;	99US-0125974.
PR	30-JUL-1999;	99US-0364862.
XX		
PA	(AVIG-) AVIGEN INC.	
PI	Couto LB, Colosi PC;	
XX		
DR	WPI; 2000-339536/29.	
XX		
PT	New recombinant adenovirus-associated vector, useful for gene therapy	











QY	195	tactgctccgcgccttcttcccgaggctgggggaggaaccgtatataaagctgactgccc	254
Db	195	caacggcctgcgtctctgccaagtgcttctgtagcaaaccccaactgctgggttgcc	254
QY	255	gtbaagttctctttcttcgcaacgggtttgcgcgcccggcaggttaagctgcaggaaatc	314
Db	255	atagccatcagcgcatgcgcgactccaaagtgtgttttgcaagagggaagcaaaaagcctc	314
QY	315	tgtgtcttaataacatcgctccaggaagatgtcttcttcttaataacatctactga	370
Db	315	ccaccgaagcctgagatgtttccacccaatgttcgaagcgtgtgttttcttgaagagaagc	374
QY	371	caactgaatcacactttcttcttcttccacagatcatcgatccaccalgcacaataagcctc	430
Db	375	aaaaagccctccaccagagccttgagctcggagagcttcgcaccacacatgcacaataagcctc	434
QY	431	tcacactgctcttctctctgctcttcttgagatctgcgtcttgaagtcgacacgaagatctac	490
Db	435	tcacactgctcttctctctgctcttcttgagatctgcgtcttgaagtcgacacgaagatctac	494
QY	491	ctgggtctgaatggaactgtcatatgagactatagcaaatgatatctcggtgcagctcttg	550
Db	495	ctgggtctgaatggaactgtcatatgagactatagcaaatgatatctcggtgcagctcttg	554
QY	551	gacgcgaagatcttccctccagagtgccaaaatctttccattcaaacctcagtcgtgtac	610
Db	555	gacgcgaagatcttccctccagagtgccaaaatctttccattcaaacctcagtcgtgtac	614
QY	611	aaaaagactcgtctgtatgaattacagagatcaccttctcaacatctcgaagcaagcca	670
Db	615	aaaaagactcgtctgtatgaattacagagatcaccttctcaacatctcgaagcaagcca	674
QY	671	ccctggatgggtctgcctcaggtgcttccacatcccaagctcgaagtcttaagatacagtgctat	730
Db	675	ccctggatgggtctgcctcaggtgcttccacatcccaagctcgaagtcttaagatacagtgctat	734
QY	731	aacacttaagaacatggtctccacatcctgtacgtcttccatgactgtgtgttatctactagg	790
Db	735	aacacttaagaacatggtctccacatcctgtacgtcttccatgactgtgtgttatctactagg	794
QY	791	aaagctctctgaagggatctgaatatgatcatcagaccagctcaaaagggagaaagaatgat	850
Db	795	aaagctctctgaagggatctgaatatgatcatcagaccagctcaaaagggagaaagaatgat	854
QY	851	aaagtcttccctgttggaagccataataatgtrctgcaagtccctgaagaagaaatgtcca	910
Db	855	aaagtcttccctgttggaagccataataatgtrctgcaagtccctgaagaagaaatgtcca	914
QY	911	atggcctctgaccacatgtgctcttaactaatatcttctcgaatgtagactgtgtaaa	970
Db	915	atggcctctgaccacatgtgctcttaactaatatcttctcgaatgtagactgtgtaaa	974
QY	971	gacttgaattcaggtctcatatgagacccta	1000
Db	975	gacttgaattcaggtctcatatgagacccta	1004

RESULT	6
AAA49232	
ID	AAA49232 standard; DNA; 12445 BP.

AC AAAA49232;

DT 26-SEP-2000 (first entry)

Vector HsQREneo for transforming endothelial cells.

KM Hemostatic; antihaemic; antidibetic; neutrophic; vector;  
KM osteopathic; anticicling; immunostimulant; gene therapy; collage;  
KM endothelial cell; peripheral blood; buffy coat cell; VEGF; PCR primer;  
vascular endothelial growth factor; bovine brain extract; haemophilia;  
Factor VIII; human; adenosine deaminase deficiency; ss;

KW sickle cell anemia; thalassemia; diabetes; alpha-fetoprotein deficiency  
 KM Alzheimer's disease; brain disease; heart disease; immune system defect;  
 KN bone fracture; osteoporosis.  
 XX  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 XX  
 W0200032750-A1.

PD 08-JUN-2000.

PF 24-NOV-1999; 99WO-US28033.

PR 24-NOV-1998; 98US-0109687.

PA (MINU ) UNIV MINNESOTA.

PA (HEBB/) HEBBEL R P.

PA (LOLL/) LOLLAR J S.

PI Hebbel RP, Lin Y, Lollar JS;  
xy

DR WPI; 2000-412303/35.  
YY

PT Expanding population of endothelial cells useful to biocompatibilize  
implantable medical devices

collagen I coated surface in culture medium comprising vascular endothelial growth factor

XX  
PS  
C]ajm 19: Fida 4: 5300: Eccl:ch

The invention relates to a method for

endothelial cells (EC) obtained from peripheral blood, by culturing, in contact with a collagen I coated surface [Buehler et al., 1994].

containing vascular endothelial growth factor (VEGF) and free of humoral

Brain extract: EC are useful for treating hemophilia by introducing it into the blood stream of a mammal so that an effective amount of

This sequence represents the vector HSOBNeo, where the human factor

sequence (H50) has been inserted. is subcloned Transgenic EC transduces

for diagnosing clotting disorders where indication or disease is

associated with a reduction in the activity of an enzyme. EC is also useful in gene therapy for treating the variety of diseases including

hemophilia, diabetes, sickle cell anemia, thalassemia, alpha-antitrypsin deficiency, brain disorders

CC for repairing bone fractures and to treat or prevent osteoporosis

Sequence 12445 BP; 3255 A; 2989 C; 2951 G; 3250 T; 0 other:

Best Local Similarity 97.1%; Pred. No. 8.6e-162;  $\chi^2$  12445;

```

0; Gbps
0; Standard
0; Maximum
0; Category

```

436

130

496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051 1052 1053 1054 1055 1056 1057 1058 1059 1060 1061 1062 1063 1064 1065 1066 1067 1068 1069 1070 1071 1072 1073 1074 1075 1076 1077 1078 1079 1080 1081 1082 1083 1084 1085 1086 1087 1088 1089 1090 1091 1092 1093 1094 1095 1096 1097 1098 1099 1100 1101 1102 1103 1104 1105 1106 1107 1108 1109 1110 1111 1112 1113 1114 1115 1116 1117 1118 1119 1120 1121 1122 1123 1124 1125 1126 1127 1128 1129 1130 1131 1132 1133 1134 1135 1136 1137 1138 1139 1140 1141 1142 1143 1144 1145 1146 1147 1148 1149 1150 1151 1152 1153 1154 1155 1156 1157 1158 1159 1160 1161 1162 1163 1164 1165 1166 1167 1168 1169 1170 1171 1172 1173 1174 1175 1176 1177 1178 1179 1180 1181 1182 1183 1184 1185 1186 1187 1188 1189 1190 1191 1192 1193 1194 1195 1196 1197 1198 1199 1200 1201 1202 1203 1204 1205 1206 1207 1208 1209 1210 1211 1212 1213 1214 1215 1216 1217 1218 1219 1220 1221 1222 1223 1224 1225 1226 1227 1228 1229 1230 1231 1232 1233 1234 1235 1236 1237 1238 1239 1240 1241 1242 1243 1244 1245 1246 1247 1248 1249 1250 1251 1252 1253 1254 1255 1256 1257 1258 1259 1260 1261 1262 1263 1264 1265 1266 1267 1268 1269 1270 1271 1272 1273 1274 1275 1276 1277 1278 1279 1280 1281 1282 1283 1284 1285 1286 1287 1288 1289 1290 1291 1292 1293 1294 1295 1296 1297 1298 1299 1300 1301 1302 1303 1304 1305 1306 1307 1308 1309 1310 1311 1312 1313 1314 1315 1316 1317 1318 1319 1320 1321 1322 1323 1324 1325 1326 1327 1328 1329 1330 1331 1332 1333 1334 1335 1336 1337 1338 1339 1340 1341 1342 1343 1344 1345 1346 1347 1348 1349 1350 1351 1352 1353 1354 1355 1356 1357 1358 1359 1360 1361 1362 1363 1364 1365 1366 1367 1368 1369 1370 1371 1372 1373 1374 1375 1376 1377 1378 1379 1380 1381 1382 1383 1384 1385 1386 1387 1388 1389 1390 1391 1392 1393 1394 1395 1396 1397 1398 1399 1400 1401 1402 1403 1404 1405 1406 1407 1408 1409 1410 1411 1412 1413 1414 141

[illegible][illegible]

210

[illegible]



Human factor VIII beta-domain deleted SQN deletion cDNA sequence

KM replication defective; recombinant retrovirus; RV; therapeutic protein  
 KM haemophilia; thrombosis; hypercoagulable disorder; liver disease; human;  
 KM hepatitis; thalassaemia; phenylketonuria; Lesch-Nyhan syndrome; diabetes;  
 KM cystic fibrosis; Duchenne's muscular dystrophy; hypercholesterolaemia;  
 KM hypoparathyroidism; adenine deaminase deficiency; HIV infection; anaemia;  
 KM Gaucher's syndrome; high blood pressure; Alzheimer's disease; autoimmune;  
 KM inflammatory disease; factor VIII; ss.

W09800541-A2

02-JUL-1997

13-AUG-1996;

(CHIR ) CHIRON CORP.

Allen JR, Barber JR, Boder M, Chang SMW, Chong K:  
De La Vega D, Depoloni J, Greengard J, Hsu DC, Ibanez CE,  
Jolly DJ, Lee R, Mittelstaedt DM, Prussak CE, Respass JG  
MPI: 1998-086966/08.

P-PSDB; AAW46246

new replication defective recombinant retro-viruses - which can be administered to provide long term systemic expression of therapeutic protein in blood, useful in, e.g. treating hyper-coagulable disorders

Example 28; Pages 210-213; 272pp; English

This cDNA encodes the beta-domain deleted SON deletion protein of human Factor VIII. This is used in the construction of recombinant retroviral vectors expressing human factor VIII. The invention provides the preparation of replication defective recombinant retrovirus (RVV) expressing a therapeutic protein. The RVV preparation is resistant to degradation by human complement and is capable of inducing long term systemic expression of the therapeutic protein when administered intravenously to a human. The long term systemic expression results in a measurable level of the therapeutic protein being produced in the blood of the human for a period of at least 30 days after the administration of the RVV vector preparation. RVV's can be used for *in vivo* delivery of therapeutic protein to treat, e.g. haemophilia A, haemophilia B, thrombosis, hypercoagulable disorders, liver diseases such as hepatitis, disorders such as thalassemia, phenylketonuria, Lesch-Nyhan syndrome, severe combined immunodeficiency (SCID), cystic fibrosis, Duchenne's muscular dystrophy, inherited emphysema, familial hypercholesterolemia, diabetes, hypoparathyroidism, adenine deaminase deficiency, alpha-1 antitrypsin deficiency, Gaucher's syndrome, anaemia, infections such as HIV infection, high blood pressure, Alzheimer's disease, autoimmune or inflammatory disease or graft versus host disease. RVV's are capable of surviving inactivation in human serum thereby allowing efficient gene transfer over prolonged periods of time.

sequence 4832 BP; 1381 A; 1089 C; 1055 G; 1307 T; 0 other;

Query Match	59.1%	Score 590.6;	DB 19;	Length 4832;
Best Local Similarity	96.2%;	Pred. No. 3.2e-161;		
Matches 605; Conservative	0;	Mismatches 24;	Indels 0;	Gaps 0

372 actgacatccactttttctttctccacaggtatcgatccacatgcaaatagagctct 433

28 attaaccttctgctctccagttgaacatttgtagcaataagtcacgaatagagctct 87

432 ccaccgctctctctgctgaccttgcgattctgctcttagtgcaccagaatactacc 491

[illegible][illegible]

QY 552 acqcaagattctctcctaqaatgcacaaatcttttcattcaaacctcaatcatatctcc 611

Db 208 acgcaagatttcctcctaagatgccaaatctttccattcaagacacctcaatcctttaga 267

612 aaagactcgttctgtagaattcagcagatcaccttttcaacatcgctaagccaagccac 671

Db 268 aaagactcgttctgtagaatcacgatacccttccaacatcgctaacccaagccac 327

QY 672 ccgagatggtctgtctagtgctcctaacatccagctgaggttatgatatacagtggtcat 731

D6 328 cctgatatggtctctgtagtctcctacatccagcgtgagtattgatatacagtggtcatlta 387

32 cactlaagacatgctcccactctgtcagtccttcatgtctgtgtatcctactgga 791

caaccagaaacacgcccacccatcgcagtccttcatgtctgtgtgtgtatcatcctaactga 447

nb 448 aaactcttaaaccaacctgaattcctccatgacacagaaagtggagaaagatgatca 851

OY 852 aaqtctccctatgaaaccatacatatatcttcacagattactcgaagcgcgggtatctctccc 011

Db 508 aagcttcctggtggaagccatacataatgtctcgcgaagcttcctaaagaaatgdtccaa 567

912 tggcctctgaccactgtgccttaactactatatttccatgtgacctgtaaaq 971

Db 568 tggcctctgacccactgtgccttaactactatatactcttctcatgtgacctgylaaag 627

QY 972 acttgaattcagcgcctattggagcccta 1000

028 acctgaatcagcctcatgtgagcccta 656

RESULT 9  
20070000

ID	standard; DNA; 4832 BP.
AAV15338	

AAV15338;  
AC  
XX

20-JUL-1998 (first entry)

number factor VIII SQN deletion mutant DNA.

retrovirus; vector; human; ss.

SS Homo sapiens.  
SS Synthatic

XX	key	Location/Qualifiers
EH		

```

CDS      12..4445
/*tag= a

```

W09800542-A2.

08-JAN-1998.

02-JUL-1997; 97WO-US11785.

04-JUN-1997; 97US-0869309.  
03-JUL-1996; 96US-0645601.  
R

13-AUG-1996; 9605-06963381.  
X

(CHIR ) CHIRON CORP.  
 Allen JR, Barber JR, Boder M, Chang SM, Chong K;  
 De la Vega D, Depolo NJ, Greengard J, Hsu DC, Ibanez CE;  
 Jolly DJ, Mittelstaedt DM, Prussak CE, Respass JG;  
 WPI; 1998-086967/08.  
 P-PSDB; AAMW4372.

New replication defective recombinant retroviruses - which express B domain-deleted human factor VIII or human factor IX for the treatment of haemophilia

Claim 6; Page 174-175; 236pp; English.

This DNA sequence includes a coding region for the B domain deletion mutant SQN (see AAMW4372) of human factor VIII. The SQN mutant is created by fusing Ser-743 to Gln-1638 of native Factor VIII (see AAMW4373) to form a Ser-Gln-Asn (SQN) link between the A2 and A3 Factor VIII domains. When compared to plasmid-derived Factor VIII, the SQN deletion does not influence the *in vivo* pharmacokinetics, but the reduced size of the molecule appears to decrease proteolytic degradation. The invention relates to preparations of replication defective recombinant retrovirus (RV) expressing a B domain-deleted human Factor VIII protein, where the recombinant RV is capable of infecting human cells, is resistant to degradation by human complement and is capable of inducing long-term (at least 30 days and up to 6 months or longer post-injection) systemic expression of Factor VIII when administered to a haemophilia A patient.

Sequence 4832 BP; 1381 A; 1089 C; 1055 G; 1307 T; 0 other;

Query Match 59.1%; Score 590.6; DB 19; Length 4832;  
 Best Local Similarity 96.2%; Pred. No. 3.2e-161;  
 Matches 605; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

```

OY 372 actgacatcaactttttttttttccacagatcgcacccatgcaatagagctct 431
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 28 attaaccttctgtctcagtgacattgtgacaaatgacatgaagaagagctct 87
OY 432 ccccccgtctcttctgtccttctgtcgttctgtcgtcgtcgtcgtcgtcgtc 491
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 88 cccacgtctcttctgtcgtcgttctgtcgttctgtcgttctgtcgtcgtcgtc 147
OY 492 tgggtgagtgagacgtgcatgaggaactatagcaagtgatctggtgagctgctg 551
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 148 tgggtgagtgagacgtgcatgaggaactatagcaagtgatctggtgagctgctg 207
OY 552 acgcaaatcttctctcctagagtgccaaatctttccattcaaacctcgtcgtgaca 611
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 208 acgcaaatcttctctcctagagtgccaaatctttccattcaaacctcgtcgtgaca 267
OY 612 aaaagactcgtgtgtgagaaatcagcgatcactttcaacatcgtcgtcgtcgtcgtc 671
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 268 aaaagactcgtgtgtgagaaatcagcgatcactttcaacatcgtcgtcgtcgtcgtc 327
OY 672 cctgtgagtggtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 731
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 328 cctgtgagtggtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 387
OY 732 cacttaagaatacgtctccatcctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 791
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 368 cacttaagaatacgtctccatcctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 447
OY 792 aagcttctgagggagctcgaatatagtatcagaccagtcgaaggaggaagaagatgata 851
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 448 aagcttctgagggagctcgaatatagtatcagaccagtcgaaggaggaagaagatgata 507
OY 852 aagcttctcctggtgaggaagcatatgtctggtcaggtcctcgaaggaggaagaagatgata 911
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 508 aagcttctcctggtgaggaagcatatgtctggtcaggtcctcgaaggaggaagaagatgata 967
  
```

OY 912 tggcctcaccacactgtgcttaccatcatatcttctcattgtagagctgtgtaaag 971  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 568 tggcctcaccacactgtgcttaccatcatatcttctcattgtagagctgtgtaaag 627

OY 972 actgaattcagagcctcattggaagccta 1000  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 628 actgaattcagagcctcattggaagccta 656

RESULT 10  
 AAT31031  
 ID AAT31031 standard; cDNA; 8967 BP.  
 XX  
 AC AAT31031;  
 XX  
 DT 19-SEP-1996 (first entry)  
 XX  
 DE Factor-VIII full-length cDNA.  
 XX  
 KW Factor-VIII; retrovirus; vector; haemophilia A; gene therapy; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 110..7165  
 FT /tag= a  
 FT sig\_peptide 110..166  
 FT /tag= b  
 FT mat\_peptide 167..7162  
 FT /tag= c  
 XX  
 PN W09621035-A2.  
 XX  
 PD 11-JUL-1996.  
 XX  
 PF 18-DEC-1995; 95MO-US16582.  
 XX  
 PR 30-DEC-1994; 94US-0366851.  
 XX  
 PA (CHIR ) CHIRON VIAGENE INC.  
 XX  
 PI Boder M, Chang S, Chi-Tang Hsu D, De Polo NJ;  
 XX  
 DR WPI; 1996-334010/33.  
 XX  
 DR P-PSDB; AAMW0465.  
 XX  
 PT Retroviral vector directing expression of full length factor VIII  
 XX  
 PS used in the gene therapy and treatment of haemophilia A  
 XX  
 PS Claim 3; Page 58-68; 86pp; English.  
 XX  
 CC A full-length cDNA clone (AAT31031) codes for human Factor-VIII  
 XX  
 CC (AAMW0465), a trace plasma glycoprotein which acts as a cofactor in  
 XX  
 CC conjunction with Factor-Xa in the activation of Factor-X.  
 XX  
 CC Retroviral vectors comprising the full-length cDNA can be efficiently  
 XX  
 CC transduced into infectious retroviral particles. These may be used to  
 XX  
 CC package cells either *in vivo* or *ex vivo*. Factor-VIII expressed  
 XX  
 CC from such transduced cells will be processed and transported in a  
 XX  
 CC fashion analogous to the expression product of a normal Factor-VIII  
 XX  
 CC gene. Retroviral particles harbouring such vectors will be useful  
 XX  
 CC in the gene therapy of haemophilia A.  
 XX  
 SQ Sequence 8967 BP; 2841 A; 1898 C; 1833 G; 2395 T; 0 other;

Query Match 59.1%; Score 590.6; DB 17; Length 8967;  
 Best Local Similarity 96.2%; Pred. No. 4.2e-161;  
 Matches 605; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

```

OY 372 actgacatcaactttttttttttccacagatcgcacccatgcaatagagctct 431
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 66 attaaccttctgtctcagtgacattgtgacaaatgacatgaagaagagctct 125
  
```





Db 666 actgaattcagccctcattgagcccta 694

## RESULT 12

ID AAN50054 standard; DNA; 8975 bp.

AC AAN50054;

DT 27-OCT-1991 (first entry)

DE Human factor VIII gene.

KW Human factor VIII; ss cDNA;

Key Location/Qualifiers

FT CDS 110..7165

PN EPI60457-A.

PD 06-NOV-1985.

PF 18-APR-1985; 85EP-0302734.

PR 20-APR-1984; 84US-0602312.

PA (GETH ) GENENTECH INC.

PI Capon DJ, Vohar GA, Lawn RM, Wood WI;

DR WPI: 1985-277976/45.

PT P-PSDB; AAP50059.

XX New recombinant functional human factor VIII or deriv. - useful for treating haemophilia and obld. as pure prod. by recombinant DNA technology.

PS Disclosure; Fig. 10A(I)-10C(III); 109pp; English.

XX The gene encodes human factor VIII. The factor VII is produced in pure form and in useful amts. using recombinant DNA technology. Factor VIII can be used to correct factor VIII deficient plasma, and activates factor X to Xa in the presence of factor IXa, Ca and phospholipid. These to activities are inactivated by antibodies specific for factor VIII. The activity of the prod. is bound to an immobilised monoclonal antibody specific for factor VIII. Factor VIII activity is activated by thrombin. The activity is bound to, and can be eluted from, immobilised von Willebrand factor. Dose of factor VIII is 20-40 units/Kg over 8 hours i.v. for maintenance therapy for haemophilia, 40 units/Kg for preoperative conditions, or 15-20 units/Kg for minor haemorrhaging.

XX Sequence 8975 BP; 2619 A; 1898 C; 1833 G; 2395 T; 0 other;

Query Match 59.18; Score 590.6; DB 6; Length 8975;

Best Local Similarity 96.28; Pred. No. 4.2e-161; Mismatches 24; Indels 0; Gaps 0;

Matches 605; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 372 actgacacacattttcttctccagagtgatccacatgcaatagagctc 431

DB 66 attaaccttggctgcagtgacatttggtagcataagtcacatgagagctc 125

QY 432 ccacgtctcttctgctgcttggcagatcgctttagtgccacagagatactac 491

DB 126 ccactgtctcttctgctgcttggcagatcgctttagtgccacagagatactac 185

QY 492 tgggtgacgtgacatgcatgagataatgacaaagtacatcggtgagctgctg 551

DB 186 tgggtgacgtgacatgcatgagataatgacaaagtacatcggtgagctgctg 245

QY 552 acgcaagatttctcctcagagtgccaaatctttccattcaacaccctcgtgtaca 611

DB 246 acgaagatttctcctcagagtgccaaatctttccattcaacaccctcgtgtaca 305

QY 612 aaaagactctgtttagaattcacgatacactttcaacatcgctaaagccac 671

DB 306 aaaagactctgtttagaattcacgatacactttcaacatcgctaaagccac 365

QY 672 cctgagtggtctgtcaggttccatccacagagtggtttatgataagtggtacata 731

DB 366 cctgagtggtctgtcaggttccatccacagagtggtttatgataagtggtacata 425

QY 732 cacttaagaacatggttccatccatctgcaagtgcttcatgtgtgtgtacccactgga 791

DB 426 cacttaagaacatggttccatccatctgcaagtgcttcatgtgtgtgtacccactgga 485

QY 792 aagctctgaggaagctgaatatgatacagaccagtcacaaaggaagaagatgata 851

DB 486 aagctctgaggaagctgaatatgatacagaccagtcacaaaggaagaagatgata 545

QY 852 aagctctcctggtgtgaagccatataatgtctgcaagtgcttcaagaagatgtccaa 911

DB 546 aagctctcctggtgtgaagccatataatgtctgcaagtgcttcaagaagatgtccaa 605

QY 912 tggcctctgacccactgtgcttactactatcttctcatgtgtgactgtgtaaaag 971

DB 606 tggcctctgacccactgtgcttactactatcttctcatgtgtgactgtgtaaaag 665

QY 972 actgaattcaagcctcattgagcccta 1000

DB 666 actgaattcaagcctcattgagcccta 694

## RESULT 13

ID AAZ38604 standard; cDNA; 8975 BP.

XX AAZ38604;

DT 28-FEB-2000 (first entry)

XX Human full-length factor VIII cDNA.

XX Factor VIII; haemophilia; proteolysis; heavy chain; light chain;

KW secretion; von Willebrand Factor; vWF; C2 domain; Intron 22; inversion;

XX non-functional; bleeding disorder; coagulation; treatment; ds.

XX Homo sapiens.

OS Homo sapiens.

Key Location/Qualifiers

FT CDS 110..7165

FT mat\_peptide 167..5233

FT signal\_peptide 167..5233

FT mat\_peptide 167..5233

FT polyA\_signal 8948..8953

FT polyA\_signal 8948..8953

PD W09959622-A1.

PD 25-NOV-1999.

PF 17-MAY-1999; 99WO-US10872.

PR 20-MAY-1998; 98US-0082000.

PA (GETH ) GENENTECH INC.





Db 502 cacttaagaacatggtccatccctgcaagtcctcatgctgtgtgtatccctactgga 561  
QY 792 aagcttctgaggagagctgaaatgatgatgatcagaccagtcacaaggagaaagaatgata 851  
|||||  
Db 562 aagcttctgaggagagctgaaatgatgatgatcagaccagtcacaaggagaaagaatgata 621  
QY 852 aagcttctgaggagagctgaaatgatgatgatcagaccagtcacaaggagaaagaatgata 911  
|||||  
Db 622 aagcttctgaggagagctgaaatgatgatgatcagaccagtcacaaggagaaagaatgata 681  
QY 912 tggcctctgaccacactgtgcttactactacatattcttcatgtgagacctgtaaaag 971  
|||||  
Db 682 tggcctctgaccacactgtgcttactactacatattcttcatgtgagacctgtaaaag 741  
QY 972 acttgaattcaggcccatctgagagcccta 1000  
|||||  
Db 742 acttgaattcaggcccatctgagagcccta 770

Search completed: January 8, 2002, 17:31:47  
Job time: 12887 sec

[illegible]

```

Dd 241 gtcgagtagtgcgcgtgaaacgttcttcttcgcacaagggtttgcgcgcgcgcgcgcgttaa 300
Qy 301 gtgcacggagatgtttgttcttaataacacacgtccacggagatgtttgttcttaataac 360
Dd 301 gtgcacggagatgtttgttcttaataacacacgtccacggagatgtttgttcttaataac 360
Qy 361 catctactgaacatcgacaatccacatttttttttttttccaaaggtatcgatccacacgtga 420
Dd 361 catctactgaacatcgacaatccacatttttttttttccaaaggtatcgatccacacgtga 420
Qy 421 aataagctctcccccgcgtctctcttcgtctgtcttttgcatttgcattagttgcacacag 480
Dd 421 aataagctctcccccgcgtctctcttcgtctgtcttttgcatttgcattagttgcacacag 480
Qy 481 aagatactacctggtgtgcagtgtgaactgttcatatggagactatataatgcagaatctgtgtga 540
Dd 481 aagatactacctggtgtgcagtgtgaactgttcatatggagactatataatgcagaatctgtgtga 540
Qy 541 gctgcctctgtgacgacaagatcttctcctctagaatgcacaatctttccattcaaacacccc 600
Dd 541 gctgcctctgtgacgacaagatcttctcctctagaatgcacaatctttccattcaaacacccc 600
Qy 601 agtcgtgtacaaaagaacctgtttgttagaattcaacgataccacttttcaaacatgcctaa 660
Dd 601 agtcgtgtacaaaagaacctgtttgttagaattcaacgataccacttttcaaacatgcctaa 660
Qy 661 gccaaagccaccctcgatgggttcgtcagggtcctacacatccacgcgcgcgcgcgttatataac 720
Dd 661 gccaaagccaccctcgatgggttcgtcagggtcctacacatccacgcgcgcgcgcgttatataac 720
Qy 721 agtgtcatltaaccttaagaacaatgtgcttccacacccgttcagttcatttcaatgtgtgtgt 780
Dd 721 agtgtcatltaaccttaagaacaatgtgcttccacacccgttcagttcatttcaatgtgtgtgt 780
Qy 781 atccctactcggaagaacctctctctgtaggagcttgatataatgatgcacaacgaatgaaggagaa 840
Dd 781 atccctactcggaagaacctctctctgtaggagcttgatataatgatgcacaacgaatgaaggagaa 840
Qy 841 aagaatgatataaagtctctccctgtgtggaagcacaatacatatgtctgtgcaggtccctgaaaga 900
Dd 841 aagaatgatataaagtctctccctgtgtggaagcacaatacatatgtctgtgcaggtccctgaaaga 900
Qy 901 gaatgttccaatgtgcctctgtgacccaactgtgccttaactactcaatcttcttcattgtga 960
Dd 901 gaatgttccaatgtgcctctgtgacccaactgtgccttaactactcaatcttcttcattgtga 960
Qy 961 cctgtttaaagaacttgaatcaggtccatattgagagcccta 1000
Dd 961 cctgtttaaagaacttgaatcaggtccatattgagagcccta 1000

RESULT 2
US-09-364-862-14
: Sequence 14, Application US/09364862
: Patent No. 6221349
: GENERAL INFORMATION:
: APPLICANT: Couto, Linda B.
: APPLICANT: Colosi, Peter C.
: TITLE OF INVENTION: ADEMO-ASSOCIATED VECTORS FOR EXPRESSION OF FACTOR VIII
: TITLE OF INVENTION: BY TARGET
: TITLE OF INVENTION: CELLS
: FILE REFERENCE: AVIGEN-03743
: CURRENT APPLICATION NUMBER: US/09/364, 862
: CURRENT FILING DATE: 1999-07-30
: EARLIER APPLICATION NUMBER: 60/125, 974
: EARLIER FILING DATE: 1999-03-24
: EARLIER APPLICATION NUMBER: 60/104, 994
: EARLIER FILING DATE: 1998-10-20
: NUMBER OF SEQ ID NOS: 14
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 14
: LENGTH: 4999

```

```

; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-364-862-14

```

Query Match	100.0%;	Score 1000;	DB 4;	Length 4999;
Best Local Similarity	100.0%;	Pred. No. 1.3e-305;		
Matches 1000; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

[illegible]







QY 868 aagcatalatgctgagcaggtccctgaagaagtgtccaaatgctctgaccact 927  
DB 487 AACCCATACATATGTGTGGCAGGCTCCCTGAAGAAGATGTCCCAATGGCCCTGACCCACT 546  
QY 928 gtgccttactactatcttctcactatgagccctggtgaagaacttaattcagacct 987  
DB 547 GTGCCTTACTACTACTATCTTCTCATGTGACCTGTGTAAGAAGCTTAATTGAGCCCT 606  
QY 988 catlgagacccta 1000  
DB 607 CATTGGAGCCCTA 619  
RESULT 6  
US-08-558-107-1  
; Sequence 1, Application US/08558107  
; Patent No. 5910481  
GENERAL INFORMATION:  
APPLICANT: VOORBERG, Johannes J.  
TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/558,107  
FILING DATE: 13-NOV-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: ISACSON, John P.  
REGISTRATION NUMBER: 33,715  
REFERENCE/DOCKET NUMBER: 30472/212  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5035 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 35..5017  
US-08-558-107-1  
Query Match 59.28; Score 592.2; DB 2; Length 5035;  
Best Local Similarity 97.98; Pred. NO. 8e-177;  
Matches 600; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 568 tagatggccaaatcttctcattcaaacctcagtggtgacaaaagactctgttl 627  
DB 187 TAGAGTGGCAAAATCTTTTTCATTACACCTCACTCGTGTACAAAAGACTCTGTGT 246  
QY 628 aagaatcagagatcaccttctcaacatcgttaagccaaagccactgagatggtctgct 687  
DB 247 AGAATTCAGGAGTCAACCTTTTCAACATCGTAAAGCCAGGCCACCTGGATGGTCTCT 306  
QY 688 aggtccctaacatccagcgtgaggttataatagatggtgtaattacacttaagaacatggc 747  
DB 307 AGGTCCATACATCCAGGCTGAGGCTTATGATACAGTGTGATACCTTAAGAACATGGC 366  
QY 748 tccccatctcagctcttctcactgctgtgtgtatccactggaagctctgagggagc 807  
DB 367 TTCCCATCTCTGTCAGCTTTTCATAGCTGTGTGTATCTACTGGAAGCTTCTGAGGAGC 426  
QY 808 tgaatatgatagtcagacagtcgaaggaggaagaagatgataaagcttccctgtgtg 867  
DB 427 TGAATATGATGATCAGACAGTCAAGGAGGAAGAGATGATTAAGTCTTCCCTGTGG 486  
QY 868 aagcatalatgctgagcaggtccctgaagaagtgtccaaatgctctgaccact 927  
DB 487 AACCCATACATATGTGTGGCAGGCTCCCTGAAGAAGATGTCCCAATGGCCCTGACCCACT 546  
QY 928 gtgccttactactatcttctcactatgagccctggtgaagaacttaattcagacct 987  
DB 547 GTGCCTTACTACTACTATCTTCTCATGTGACCTGTGTAAGAAGCTTAATTGAGCCCT 606  
QY 988 catlgagacccta 1000  
DB 607 CATTGGAGCCCTA 619  
RESULT 7  
US-09-243-539-1  
; Sequence 1, Application US/09243539  
; Patent No. 6130203  
GENERAL INFORMATION:  
APPLICANT: VOORBERG, Johannes J.  
TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/243,539  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/558,107  
FILING DATE: 13-NOV-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: ISACSON, John P.  
REGISTRATION NUMBER: 33,715  
REFERENCE/DOCKET NUMBER: 30472/212  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5035 base pairs  
TYPE: nucleic acid







372 actgacatccacttttcttcttctcaccaggtatcgatccacatgcaatagagctct 431  
107 ATTAACTTTTGGCTTCTCCAGTTGACATTTGTACCAATAGATGCAATAGAGCTCT 166  
432 caactgtctcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 491  
167 CCACCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 226  
492 tgggtgacgtggaactgcatgagacataatgcaagatgctcgtgagctcgtg 551  
227 TGGGTGACAGTGAAGTGCATGAGCTATATGCAAGATGATCTCGTACGCTGTGG 286  
552 acgcaagatcttctctcctcctcctcctcctcctcctcctcctcctcctcctcct 611  
287 ACGCAAGATTTCTCTCTAGAGGCAAAATTTTTCATTCACACCTCAGTCTGTACA 346  
612 aaaaagactgttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 671  
347 AAAAGACTGTTGTTGTAGATTCACGGTTCCACTTTTCAACTTCGCTTAAGCCAGCCAC 406  
672 cctgagatggtctgctgtaggctcctacatccaggtcaggttataatagatgctat 731  
407 CCGATGAGGTCTGCTAGTGTCTTACATCCAGCTGAGGTTATGATACAGTGTCTATTA 466  
732 caattaagaacatgcttctcctcctcctcctcctcctcctcctcctcctcctcct 791  
467 CACTTAAGAACATGCTTCCATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 526  
792 aagctctcaggaagcgtatgatatgatatgatatgatatgatatgatatgatatgata 851  
527 AAGCTTCTGAGGAGCTGATATGATGATGATGATGATGATGATGATGATGATGATGAT 586  
852 aagctctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 911  
587 AAGCTTCTGAGGAGCTGATATGATGATGATGATGATGATGATGATGATGATGATGAT 646  
912 tggcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 971  
647 TGGCCTCTGACCCACGCTGCTTACTCTACTCTACTCTACTCTACTCTACTCTACTCTA 706  
972 actgaatcagcgcctcctcctcctcctcctcctcctcctcctcctcctcctcct 1000  
707 ACTTGAATTCAGGCTCTCATTTGAGGCCCTTA 735

RESULT 12  
US-08-474-503-1  
Patent No. 5744446  
GENERAL INFORMATION:  
APPLICANT: Emory University  
TITLE OF INVENTION: Hybrid: Human/Animal Factor VIII  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Kilpatrick & Cody  
STREET: 1100 Peachtree Street, Suite 2800  
CITY: Atlanta  
STATE: Georgia  
COUNTRY: US  
ZIP: 30309  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/474,503  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Pratt, John S.  
REGISTRATION NUMBER: 29,476  
REFERENCE/DOCKET NUMBER: EMU106CIP(3)

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 404-815-6500  
TELEFAX: 404-815-6555  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9009 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHEetical: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Homo sapien  
TISSUE TYPE: Liver  
FEATURE:  
NAME/KEY: misc\_feature (Domain Structure)  
LOCATION: 5125..7053  
OTHER INFORMATION: /note= "Equivalent to the A3-C1-C2  
OTHER INFORMATION: domain"  
FEATURE:  
NAME/KEY: misc\_feature (Domain Structure)  
LOCATION: 1..227  
OTHER INFORMATION: /note= "Equivalent to the A1-A2 domain."  
FEATURE:  
NAME/KEY: Domain  
LOCATION: 1..227  
OTHER INFORMATION: /note= "cDNA encoding human factor  
OTHER INFORMATION: VIII."  
US-08-474-503-1

Query Match 58.9%; Score 589; DB 1; Length 9009;  
Best Local Similarity 96.0%; Pred. No. 1,1e-175;  
Matches 604; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

372 actgacatccacttttcttcttctcaccaggtatcgatccacatgcaatagagctct 431  
107 ATTAACTTTTGGCTTCTCCAGTTGACATTTGTACCAATAGATGCAATAGAGCTCT 166  
432 caactgtctcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 491  
167 CCACCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 226  
492 tgggtgacgtggaactgcatgagacataatgcaagatgctcgtgagctcgtg 551  
227 TGGGTGACAGTGAAGTGCATGAGCTATATGCAAGATGATCTCGTACGCTGTGG 286  
552 acgcaagatcttctctcctcctcctcctcctcctcctcctcctcctcctcctcct 611  
287 ACGCAAGATTTCTCTCTAGAGGCAAAATTTTTCATTCACACCTCAGTCTGTACA 346  
612 aaaaagactgttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 671  
347 AAAAGACTGTTGTTGTAGATTCACGGTTCCACTTTTCAACTTCGCTTAAGCCAGCCAC 406  
672 cctgagatggtctgctgtaggctcctacatccaggtcaggttataatagatgctat 731  
407 CCGATGAGGTCTGCTAGTGTCTTACATCCAGCTGAGGTTATGATACAGTGTCTATTA 466  
732 caattaagaacatgcttctcctcctcctcctcctcctcctcctcctcctcctcct 791  
467 CACTTAAGAACATGCTTCCATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 526  
792 aagctctcaggaagcgtatgatatgatatgatatgatatgatatgatatgatatgata 851  
527 AAGCTTCTGAGGAGCTGATATGATGATGATGATGATGATGATGATGATGATGATGAT 586  
852 aagctctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 911  
587 AAGCTTCTGAGGAGCTGATATGATGATGATGATGATGATGATGATGATGATGATGAT 646  
912 tggcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 971

|||||  
Db 647 TGGCCTTGACCCACTGTGCTTACCTACATATCTTCTCATGTGACCTGTAAAG 706  
Qy 972 actgaatcagccctcatctgagacccta 1000  
|||||  
Db 707 ACTGATTCAGGCGCTCATGTGAGCCCTA 735

RESULT 13  
US-08-670-707A-1  
; Sequence 1, Application us/08670707A  
; Patent No. 5859204  
GENERAL INFORMATION:  
APPLICANT: Lollar, John S.  
TITLE OF INVENTION: Hybrid Human/Animal Factor VIII  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.  
STREET: 5370 Manhattan Circle Suite 201  
CITY: Boulder  
STATE: Colorado  
COUNTRY: USA  
ZIP: 80303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patenting Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US 08/670,707A  
FILING DATE: 26-JUN-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US94/13200  
FILING DATE: 15-NOV-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/212,133  
FILING DATE: 11-MAR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/864,004  
FILING DATE: 07-APR-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Greenlee, Lorraine L.  
REGISTRATION NUMBER: 27,894  
TELEPHONE: 303/499-8089  
TELEFAX: 303/499-8080  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9009 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
MOLECULE TYPE: not relevant  
TOPOLOGY: not relevant  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
TISSUE TYPE: Liver  
FEATURE:  
NAME/KEY: misc.feature  
LOCATION: 1..2277  
OTHER INFORMATION: /product="Domain Structure"  
OTHER INFORMATION: /note="Equivalent to the A3-C1-C2 domain"  
FEATURE:  
NAME/KEY: misc.feature  
LOCATION: 1..2277  
OTHER INFORMATION: /product="Domain Structure"  
OTHER INFORMATION: /note="Equivalent to the A1-A2 domain"  
FEATURE:  
NAME/KEY: misc.feature  
LOCATION: 1..2277

OTHER INFORMATION: /product="Domain"  
OTHER INFORMATION: /note="cDNA encoding human factor-VIII"  
US-08-670-707A-1

Query Match 58.9%; Score 589; DB 2; Length 9009;  
Best Local Similarity 96.0%; Pred. No. 1,1e-175;  
Matches 604; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

Qy 372 actgacatccactcttcttctccacagatcagatccacatgaaatagagctc 431  
Db 107 ATTAACTTTCTCTCTCCAGTTGAACATTTGTACCAATAGCATGCAATAGAGCTCT 166  
Qy 432 ccactgctcttctctgaccttgcgactctgcttgcgacctgagccaggaatactacc 491  
Db 167 CCACCTGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 226  
Qy 492 tgggtgcagtggaactgcaatgagactatataagaaagatcgtgagctgctg 551  
Db 227 TGGGTGACAGTGGAAGTGCATGAGGACATATGCAAAAGATCTGCTGCTGCTGCTG 286  
Qy 552 acgcaagattctctctagatgagtcgaatcttccatcaacactcagtgatga 611  
Db 287 ACGCAAGATTCTCTCTAGAGTGCCTCAATCTTTCCATTCACACCTCATCTCTGACA 346  
Qy 612 aaaaagcctcttgtagaatcaagatcaacattcaacatcgaagcgaagccac 671  
Db 347 AAAAAGCTCTGTTGTAAATTAACAGGTTCAACCTTTCAACATGCTTAAGCAGGCCAC 406  
Qy 672 cctggaatgctgctagtgctccacacacagcagtgagttatgatacagtgatca 731  
Db 407 CCTGATGGGTCTGCTAGAGTGCCTCAATCTTTCCATTCACACCTCATCTCTGACA 466  
Qy 722 caactaagaacatgagctccacatcctgctgctgctgctgctgctgctgctgctg 791  
Db 467 CACTTAAGAACATGAGCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 526  
Qy 792 aagctctgagagagctgataatgataatcagacagtcgaagagagagagagatgata 851  
Db 527 AAGCTTTCAGAGGAGCTGAATATGATGATGATGATGATGATGATGATGATGATGATG 586  
Qy 852 aagctctgatgata 911  
Db 587 AAGCTTTCAGAGGAGCTGAATATGATGATGATGATGATGATGATGATGATGATGATG 646  
Qy 912 tggcctcagaccactgacctaccatcactcactcactcactcactcactcactcactc 971  
Db 647 TGGCCTTGACCCACTGTGCTTACCTACATATCTTCTCATGTGACCTGTAAAG 706  
Qy 972 actgaatcagccctcatctgagacccta 1000  
Db 707 ACTGATTCAGGCGCTCATGTGAGCCCTA 735

RESULT 14  
US-09-037-601-1  
; Sequence 1, Application us/09037601  
; Patent No. 6180371  
GENERAL INFORMATION:  
APPLICANT: Lollar, John S.  
TITLE OF INVENTION: Hybrid Human/Animal Factor VIII  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.  
STREET: 5370 Manhattan Circle Suite 201  
CITY: Boulder  
STATE: Colorado  
COUNTRY: USA  
ZIP: 80303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS

```

SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/037,601
FILING DATE: 26-JUN-1996
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US94/13200
FILING DATE: 15-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/212,133
FILING DATE: 11-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/864,004
FILING DATE: 07-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Feilber, Donna M.
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 75-95F
TELECOMMUNICATION INFORMATION:
TELEPHONE: 303/499-8080
TELEFAX: 303/499-8089
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 9009 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: not relevant
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
TISSUE TYPE: Liver
FEATURE:
NAME/KEY: misc-feature
LOCATION: 5125..7053
OTHER INFORMATION: /product= "Domain Structure"
OTHER INFORMATION: /note= "Equivalent to the A3-C1-C2 domain"
FEATURE:
NAME/KEY: misc-feature
LOCATION: 1..2277
OTHER INFORMATION: /product= "Domain Structure"
OTHER INFORMATION: /note= "Equivalent to the A1-A2 domain"
FEATURE:
NAME/KEY: misc-feature
LOCATION: 1..2277
OTHER INFORMATION: /product= "Domain"
OTHER INFORMATION: /note= "cDNA encoding human factorVIII"
US-09-037-601-1

```

Query Match 58:9% Score 589; DB 4; Length 9009;  
 Best Local Similarity 96.0%; Pred. No. 1,1e-175;  
 Matches 604; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

```

QY 372 actgacatcacttttcttctccacaggtacgataccacatgcaataagagctct 431
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 107 ATTAACTTTTCTCTCTCAGTTGAACATTTGTACATAGTCTATGCAATAGAGCTCT 166
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 432 ccactgtcttctctgtgacctttgagattctgctttagtgcacacagaatactacc 491
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 167 CCACCTGCTTCTTCTGCGCTTTGCGATCTCTGTTAGTCCACGAGAGACTACTACC 226
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 492 tgggtgcagtgagactgtcatgagactatatagcagaagatctggtgagctgtg 551
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 227 TGGGTGCAGTGAACATGCTATGGAAGCAAGATGATCTCGGTAGCTGCTGTGG 286
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 552 acgcaagattctctctgtagagtgccaatacttccattcaacacccagtcgtgtaca 611
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 287 ACGCAAGATTCTCTCTAGAGTGCAGAAATCTTTCCATTCAACACCTCAGTCGTGTACA 346
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 612 aaaaagcttgtttagaatcacgagatccacttcttaacacgcgtaagccaagccac 671
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

```

```

Db 347 AAAAGACTGTTGTAGAAATTCACGGTTCACCTTTTCAACATCGCTAAGCAGGCCAC 406
QY 672 ccttgatgggtctgtctgagttccatccacaggtgaggttttataagatggtacata 731
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 407 CCTGATGGGTCTCTAGTCTTACCATTCAGAGCTTGAGTTTATGATGAGTGTCTATTA 466
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 732 cacltaagaacatgcttccatctctgacgtcttcatagtctgtgtgtatccactgta 791
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 467 CACTTAAGAACATGGCTCCATCCATCCATGTCAGTCTTATGCTGTGGTATCTACTAGGA 526
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 792 aagcttctgaggagctgtaataatgataagacagtcacaaggaggaagaagatgata 851
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 527 AAGCTTCTGAGGGAGCTGAATATGATGATGATGATGATGATGATGATGATGATGATA 586
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 852 aagcttcccttggtggaagccatcatatgtctggagagtcctcgaagaagaatggtccaa 911
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 587 AAGCTTCTCCCTGGTGGAGAGCCATATATGTGTGGAGTGTGTAAGAGAAAGTGTCAA 646
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 912 tggccttgacccactgtgcttaccactacatacttctcatgtgtgagctgtgtaaaag 971
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 647 TGGCTCTGAGCCACCTGCTTACCTTACCTATATCTTTCTCATGTGACCTGTGTAAG 706
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 972 actgaaatcaggcctcattggagcccta 1000
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 707 ACTGAATTCAGGCTCATTTGAGACCCCTA 735
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

```

RESULT 15  
 PCT-US93-03275-3  
 Sequence 3, Application PC/RUS9303275  
 GENERAL INFORMATION:  
 APPLICANT: Lollar, John S.  
 APPLICANT: Runge, Marschall S.  
 TITLE OF INVENTION: Hybrid Human/Porcine Factor VIII  
 NUMBER OF SEQUENCES: 6  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Kilpatrick & Cody  
 STREET: 1100 Peachtree Street, Suite 2800  
 CITY: Atlanta  
 STATE: Georgia  
 COUNTRY: US  
 ZIP: 30309-4530  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: PCT/US93/03275  
 FILING DATE: 19930407  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/864004  
 FILING DATE: 07-APR-1992  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Pabst, Patrea L.  
 REGISTRATION NUMBER: 31,284  
 REFERENCE/DOCKET NUMBER: EMU 106PCT  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 404-815-6558  
 TELEFAX: 404-815-6555  
 INFORMATION FOR SEQ ID NO: 3:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 9009 base pairs  
 TYPE: NUCLEIC ACID  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 FRAGMENT TYPE: N-terminal  
 ORIGINAL SOURCE:  
 ORGANISM: Homo sapien





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 8, 2002, 15:13:30 ; Search time 4541.38 Seconds  
(without alignments)  
2366.192 Million cell updates/sec

Title: US-09-740-211-14\_COPY\_1\_1000

Perfect score: 1000  
Sequence: 1 cagccctgcagcagcagctgcg.....cagccctcatctggagcccta 1000

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 11351937 segs, 5372899281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match: 0%  
Maximum Match: 100%  
Listing first 45 summaries

Database :

EST:  
1: em\_estfun:  
2: em\_estfun:  
3: em\_estfun:  
4: em\_estfun:  
5: em\_estfun:  
6: em\_estfun:  
7: em\_estfun:  
8: em\_estfun:  
9: em\_estfun:  
10: em\_estfun:  
11: em\_estfun:  
12: em\_estfun:  
13: em\_estfun:  
14: em\_estfun:  
15: em\_estfun:  
16: em\_estfun:  
17: em\_estfun:  
18: em\_estfun:  
19: em\_estfun:  
20: em\_estfun:  
21: em\_estfun:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	323.4	32.3	954	11	BG477974
2	98.6	9.9	456	11	AA817854
3	98.6	9.9	625	10	AI526753
4	98.6	9.9	658	11	BG923886
5	98.6	9.9	783	11	BG977142
6	98.6	9.9	3739	12	AK014835
7	98.2	9.8	1006	10	AI225600
8	97.8	9.8	662	11	BF150467
9	97.8	9.7	632	11	BG662595
10	95.8	9.6	521	10	AA865338
11	95.8	9.6	878	11	BE912296
12	95.8	9.6	918	10	AL556703

13	94.2	9.4	634	10	AV653260
14	93.8	9.4	477	11	AA461838
15	91.8	9.2	576	11	BP914540
16	91.6	9.2	913	11	BE916015
17	91.6	9.2	954	11	BG172588
18	90.4	9.0	722	11	BF136574
19	89.8	9.0	936	11	BF138391
20	89.2	8.9	695	11	BF107719
21	89.2	8.9	839	10	BE371524
22	89.2	8.9	836	11	BE371524
23	88.2	8.8	880	11	BE371524
24	88.2	8.8	567	11	AA873581
25	88.2	8.8	1002	11	BF532535
26	87.8	8.8	657	11	BI428548
27	87.6	8.8	621	11	BG914512
28	86.6	8.7	472	10	BE308999
29	86.6	8.7	530	10	BE308999
30	86.6	8.7	667	10	BE285283
31	86.6	8.7	691	11	BG975622
32	86.6	8.7	796	11	BF383789
33	86.6	8.7	837	11	BF144018
34	86.6	8.7	893	10	BE283339
35	86.2	8.6	642	11	BF158434
36	86.2	8.6	843	11	BF122699
37	86.2	8.6	880	10	BE286923
38	85.4	8.5	732	11	BG568672
39	85.4	8.5	475	10	AI265337
40	85.4	8.5	670	11	BG976890
41	85.4	8.5	921	11	BF138986
42	84.6	8.5	651	11	BF156677
43	83.3	8.3	684	11	BF384924
44	82.6	8.3	509	11	BF604025
45	81.6	8.2	622	10	AW175422

#### ALIGNMENTS

RESULT 1  
LOCUS BG477974 954 bp mRNA EST 21-MAR-2001  
DEFINITION 602522966F1 NIH\_MGC\_20 Homo sapiens cDNA clone IMAGE:4641352 5',  
mRNA sequence.  
ACCESSION BG477974  
VERSION BG477974.1 GI:13410253  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 954)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
TITLE Unpublished (1999)  
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cga@bbs.fda.gov  
Tissue Procurement: ATCC/CDT/DTF  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: LINC1408 row: c column: 17  
High quality sequence stop: 626.  
Location/Qualifiers  
1. 954  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4641352"  
/clone\_lib="NIH\_MGC\_20"  
/tissue\_type="melanotic melanoma"  
/lab\_host="DH10B (phage-resistant)"

#### FEATURES

source



DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
U42e03.y1 Sugano mouse liver mlia Mus musculus cDNA clone IMAGE:1922620 5' similar to gb:M13699 CERULOPLASMIN PRECURSOR (HUMAN); mRNA sequence.	A1526753	GI:4440888								
EST.										
house mouse.										
Mus musculus										
Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi; Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Murinae: Mus.										
1 (bases 1 to 625)										
Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Stepien,M., Thistling,B., Allen,M., Bowers,Y., Peterson,B., Swaller,T., Gibbons,M., Page,D., Harvey,N., Schurr,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.										
The WashU-NCI Mouse EST Project 1999										
Unpublished (1999)										
Contact: Marra,M./washu-NCI Mouse EST Project 1999										
Washington University School of Medicine										
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA										
Tel: 314 286 1800										
Fax: 314 286 1810										
Email: mouseest@wustl.wustl.edu										
This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.										
Seq primer: custom primer used										
High quality sequence stop: 510.										
Location/Qualifiers										
1..625										
/organism="Mus musculus"										
/strain="C57BL"										
/db_xref="taxon:10090"										
/clone="IMAGE:1922620"										
/clone_lib="Sugano mouse liver mlia"										
/sex="female"										
/dev_stage="adult"										
/lab_host="DH10B"										
/note="Organ: liver; Vector: pME18S-FL3; Site_1: DraIII (CACGCTG); Site_2: DraIII (CACCAATG); 1st strand cDNA was primed with an oligo(dT) primer [ATGGCCCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [GTTGGCCCTACTG], digested and cloned into distinct draIII sites of the pME18S-FL3 vector (5' site CACTGCTG, 3' site CACCAATG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCGCTCTAAAGACTGCG and 3' end primer CGACTCGAGCTCGACACA."										
BASE COUNT	185 a	126 c	140 g	174 t						
ORIGIN										
Query Match	9.9%;	Score 98.6;	DB 10;	Length 625;						
Best Local Similarity	55.0%;	Pred. No. 8.3e-18;								
Matches 216;	Conservative 0;	Mismatches 174;	Indels 3;	Gaps 1						
OY	608	tacaaaagaagctctgttctgtagaataccaagcagcatttccaacatgcgtgaagccaag	667							
Db	228	TATTAAGAAAGCCCTTTATTTTGGTACACAGAGGACGACCTTTAGTAAAGATAGACAA	287							
OY	668	ccacccctgagtggtcgtcgtcagctccacatccacagcgtgaggttatataacagtgyc	727							
Db	288	CCAGCCTGGCTAGAGGTTTATAGCCCTGTCATCAAGAGCTAAATTGAAGATAAGTTTAT	347							
OY	728	ataacactaagaacatgctccacatccctgtcagctcctaagcgtgtgtgtatcctac	787							
Db	348	GTTACATTAAAGAACCTTCCTCTAGAGATTACACTTTTTCATGCACATGGGCTAACGTAC	407							
OY	788	tggaaagcttcctgagtgagctgataatgatcatcagaccacccaaggggagaagaagat	847							

[illegible]

Query Match	9.98;	Score 98.6;	DB 11;	Length 783;
Best Local Similarity	55.08;	Pred. No. 9e-18;		
Matches 216; Conservative	0;	Mismatches 174;	Indels 3;	Gaps 1

acaaagaagactctgttctgtaaatcacaatcacgaatacaacctttccacatcgctaagccaagg 667  
+ | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
ATAAGAAGGCGCCTTATTTTGTGATACACAGATTGGCACCTTAGTAAGACTAATAGCAAA 315  
+ | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
ccacccttgatggctgcctcgaagtcctaccacatccagagctgaagtattgatcatagtgatc 727  
+ | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
ccacccttgatggctgaagcttt 375  
+ | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
tttcaacttaagaacatggtctcccatccctgcctcagctctcagtcgctgttggtatccctac 787  
+ | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
tgaagaagctctctgagggagcttgatatgatgatgcagaccagtcacaaggctgaattaaagtttat 907  
+ | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
ataaagctctcccctggaggagacatatatgctctggcaggtcctctgtaaagaagatgtg 907  
+ | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
ccaccaaatagctgtcccccgacacacagatgctgtagtgtgctcattgcc--aatagcccaact 552  
+ | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
caatggcctctgagcccaactgtgcttacctactactatctcttctcattgtgagcttgta 967  
+ | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
cttggaagagggagacagcaattgtgtgacacagattttaccactccactccatgttgatgtcca 612  
+ | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
aagacttgaaattcaagcctcatgtggagcccta 1000  
+ | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
AAAGATATTGCATCAGCAGCATGAGCACTCTGA 645

AK014835 3739 bp mRNA HTC 05-JUL-2001  
Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:A92J507J22, full insert sequence.  
AK014835  
AK014835.1 GI:12852907  
CAP trapper.  
Mus musculus (strain:C57BL/6J) adult male testis cDNA to mRNA, clone\_lib:RIKEN full-length enriched mouse cDNA library  
clone:A92J507J22.

Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 3739)  
Carinci,P. and Hayashizaki,Y.  
High-efficiency full-length cDNA cloning methods in enzymology. 303, 19-44 (1999)  
99279253  
10349636

2 (bases 1 to 3739)  
Carinci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
Genome research. 10 (10), 1617-1630 (2000)  
20499374  
99279253

3 (bases 1 to 3739)  
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carinci,P., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishino,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Matshiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsunura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kita,A. and Hayashizaki,Y.  
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer  
Genome research. 10 (11), 1757-1771 (2000)  
20530913  
20530913  
11076861

4 (bases 1 to 3739)  
The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.





## source

1. 632  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /clone\_lib="IMAGE:4917258"  
 /clone\_lib="NIH\_CGAP\_Mam4"  
 /tissue\_type="tumor, gross tissue"  
 /lab\_host="DH10B"  
 /note="Organ: mammary; Vector: pcwv-sport6; Site: 1: NotI; Site: 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert 2.5 kb. Library constructed by Life Technologies, catalog # 12018-016. Investigators providing samples: Lothar Hennigshausen/Priscilla Furth, NIH Reference for transgenic model: Li et al., Cell Growth and Differentiation 7, 3-11 (1996). Note: this is a NCI\_CGAP Library."

BASE COUNT 195 a 125 c 139 g 173 t  
 ORIGIN

Query Match 9.7%; Score 97; DB 11; Length 632;  
 Best Local Similarity 54.7%; Pred. No. 2.5e-17;  
 Matches 215; Conservative 0; Mismatches 175; Indels 3; Gaps 1;

608 tacaaaagactctgttgaatcagcagcatttcaaatcgttaagcgaagg 667  
 210 TATATAGAGGCCCTTATTTGATGACAGATGACACCTTAGTCAGATAGCAAA 269  
 668 ccaccctgtagtgcctgaagctcctacacagcgtgaagttatatacagtc 727  
 270 CCACCTGGCTAGAGGCTTTTAAAGCCCTGTATCAAGCTGAAGATTAAATTTAT 329  
 728 attacttaagaacatgctccacatccctgacatcctgaagctgctgctac 787  
 330 GTTCACTTAAAGAACCTTGCTCTAGATGTACACTTTTCATGACATGCGGAAGCTAC 389  
 788 tgaagaagctctcgtgaggaagctgaatgatgatcagcagcgaaggaagaag 847  
 390 ACCAAGAGATGATGAGGAGCGCTACCTGACACACCCGATATTTTACCGGCTGAT 449  
 848 gataaagctcctcgtgtgtaagcacaatataatgctcgaagctcctgaagaag 907  
 450 GACAAAGCTCTCCCGGACACAGTATGTATGTGTGATGCTGATCC---AATGAGCAAT 506  
 908 ccaatgacctcgtgacacactgctcctacatcatatcttcaatgtgagcctg 967  
 507 CCTGAGAGGAGAGACAGATTTGTGTACACAGATTACACACGCCATGTTGCTCCA 566  
 968 aaagactgaattcaagcctcatctgagcccta 1000  
 567 AAAGATATTGCATCAGACATCATAGACCTCTA 599

RESULT 10  
 AA865338 521 bp mRNA EST 29-APR-1998  
 LOCUS og88d10.s1 NCI\_CGAP\_Kid5 Homo sapiens cDNA clone IMAGE:145379 3'  
 DEFINITION similar to gb:U13699 CERULOPLASMIN PRECURSOR (HUMAN); mRNA sequence.

ACCESSION AA865338  
 VERSION AA865338.1 GI:2957614  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens

REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 521)  
 NCBI\_CGAP http://www.ncbi.nlm.nih.gov/ncigap.  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 JOURNAL Unpublished (1997)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.

Emmert-Buck, M.D., Ph.D.  
 CDNA Library Preparation: M. Bento Soares, Ph.D.  
 CDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI\_CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 www.bio.llnl.gov/bbrp/image/image.html  
 Insert Length: 656 Std Error: 0.00  
 Seq primer: -40m13 fwd. ET from Amersham  
 High quality sequence stop: 360.  
 Location/Qualifiers  
 1. 521

## FEATURES

## source

/organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="IMAGE:145379"  
 /clone\_lib="NCI\_CGAP\_Kid5"  
 /tissue\_type="2 pooled tumors (clear cell type)"  
 /lab\_host="DH10B"  
 /note="Organ: kidney; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site: 1: Not I; Site: 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer (5' AACCTGAGAAATTCGGCGCGCAATATTTTATTTTATTTT 3'), double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization. Library constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 138 a 100 c 107 g 176 t  
 ORIGIN

Query Match 9.6%; Score 95.8; DB 10; Length 521;  
 Best Local Similarity 52.7%; Pred. No. 5.2e-17;  
 Matches 208; Conservative 0; Mismatches 187; Indels 0; Gaps 0;

606 tgaataaagactctgttgaatcagcagcatttcaaatcgttaagcgaagg 665  
 425 TATATAGAGGCCCTTATTTGATGACAGATGACACCTTAGTCAGATAGCAAA 366  
 666 ggcaccctgtagtgcctgaagctcctacacagcgtgaagttatatacagtc 725  
 365 AACCGCTGTGGCTTTTAAAGCCCTGTATCAAGCTGAAGCTGAAGATTAAG 306  
 726 tcattacactaaagaatgctcctcctacatcctgacgtctcctgctgtgtg 785  
 305 ATGTACACTTAAAGACCTTGCTCTTACGCTTACACCTTTCATGACATGAAT 246  
 786 actgaaagctctgaggaagctgaatataatgatcagcagcgaaggaagaag 845  
 245 ACTATAAGAGACATGAGGGGGCCATCTACCTGATACACACAGATTTCACAA 186  
 846 atgataaagctcctcgtgaggaagcacaatataatgctgagcagcgttcctgaagaag 905  
 185 ATGACAAAGTATATTCACAGAGAGCATATATCATGTTGCTTGCCTGCTGAA 126  
 906 gtccaatgacctgagcaccactgctcctacatcatatcttctcattgtgagc 965  
 125 GTCTGGGGAAGAGATGCAATTGTGTAGCTAGGATTTACATCCCATTTGAG 66  
 966 taaagactgaattcaagcctcatctgagcccta 1000  
 65 CAAGAATATTGGCTCAGAGACTCATCGACCTTTA 31

## RESULT 11

BE912296 878 bp mRNA EST 29-SEP-2000  
 LOCUS BE912296  
 DEFINITION 601666013p1 NCI\_CGAP\_Mam1 Mus musculus cDNA clone IMAGE:3966051 5',  
 mRNA sequence.

ACCESSION BE912296  
 VERSION BE912296.1 GI:10409387  
 KEYWORDS EST.  
 SOURCE house mouse.

## ORGANISM

Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathia; Muridae; Murinae; Mus.  
REFERENCE  
1 (bases 1 to 878)  
AUTHORS  
NIH-MGC  
TITLE  
National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL  
Unpublished (1999)  
COMMENT  
Contact: Robert Stausberg, Ph.D.  
Email: c9apbs-r@mail.nih.gov  
Tissue Procurement: Gilbert Smith, Ph.D.  
CDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
plate: L1AM9138 row: j column: 04  
High quality sequence stop: 596.

## FEATURES

## Source

1..878  
/organism="Mus musculus"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:3966051"  
/clone\_lib="NCI\_CGAP\_Mam1"  
/tissue\_type="tumor, biopsy sample"  
/dev\_stage="3 months, virgin"  
/lab\_host="DH10B"  
/note="Organ: mammary; Vector: PCMV-SPORT6; Site: 1; Salt:  
Site: 2; Not: Cloned unidirectionally. Primer: Oligo dT.  
Library constructed by Life Technologies. Investigator  
providing samples: Gilbert Smith, NIH"

## BASE COUNT

268 a 188 c 195 g 227 t

## ORIGIN

Query Match 9.6%; Score 95.8; DB 11; Length 878;  
Best Local Similarity 54.8%; Pred. No. 6.2e-17;

Matches 212; Conservative 0; Mismatches 172; Indels 3; Gaps 1;

QY 608 tcaaaaagactcgttctgtagaattcaacgagtcacatttcaacatgcgaagcaag 667  
DB 181 TATTAAGAAAGCCCTTTATTTGATGACACAGATGGACACCTTAGTAACATATAGACAAA 240  
QY 668 ccacctgtagtgcctgtagtctcaccatccagagctgaggtttatgatacagtgctc 727  
DB 241 CCAGCCTGGCTAGGCTTTTATAGCCCTGTCATCAAGCTGAAGTTGAAGATTAAGTTTAT 300  
QY 728 attacacttaagaacatgctcctccatcctgtagtctcctgtagtctgtagtctcctac 787  
DB 301 GTTCACTTAAGAAGACCTTGCCTTAGATGATCTTTCATGACATGGGGTAAACGTAC 360  
QY 788 tggaaagctctgagagagctggaatagatgatacagacagcaagcaaggaagaaagaaat 847  
DB 361 ACCAAGGAGTATGAGGAGGAGCGCTACCTGACCAACACCTGATTTTCAACGGGGCTGAT 420  
QY 848 gataaagctccctgggaggaagcacaatcatatgctgagagagctcgtgaagaaagaaatg 907  
DB 421 GACAAAGTGTCCCGGACCAACAGTATGTGATGTGCTGATGCC---AATGACCAAGT 477  
QY 908 ccaatgagctctgagagagctgtagtctcctcattcattctctcattgtagtctgtagt 967  
DB 478 CCGTAGAGGAGGAGACAGCAATTGTGTGACACGAGATTACCACTCCCATGTTGATGCTCA 537  
QY 968 aaagactgtaattcaagpctcatcgga 994  
DB 538 AAAAGTATTTGCATCAGGATTCATAGGA 564

## RESULT 12

AL556703 918 bp mRNA EST 16-FEB-2001  
LOCUS AL556703 L1\_NFL006\_PL2 Homo sapiens CDNA clone CSDDK005Y102 5  
DEFINITION

prime, mRNA sequence.  
AL556703  
AL556703.1 GI:12899623  
KEYWORDS  
EST.  
SOURCE  
human.  
ORGANISM  
Homo sapiens

REFERENCE  
1 (bases 1 to 918)  
AUTHORS  
Li, W.B., Gruber, C., Jesse, J. and Polayes, D.  
TITLE  
Full-length cDNA libraries and normalization  
JOURNAL  
Unpublished (2001)  
COMMENT  
Contact: Genoscope  
Genoscope - Centre National de Sequences  
BP 191 91006 Evry cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

## FEATURES

## Source

1..918  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="CSDDK005Y102"  
/clone\_lib="L1\_NFL006\_PL2"  
/tissue\_type="placenta"  
/note="Vector: PCMVSPORT 6; Site: 1; Not: 1st strand cDNA  
was primed with a NotI-oligo(dT) primer. Five prime end  
enriched, double-stranded cDNA was digested with Not I and  
cloned into the Not I and Eco RV sites of the PCMVSPORT 6  
vector. Library was normalized. Library was constructed by  
Life Technologies. Contact : Feng Liang Life Technologies,  
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371  
Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com"

## BASE COUNT

303 a 169 c 192 g 253 t

## ORIGIN

Query Match 9.6%; Score 95.8; DB 10; Length 918;  
Best Local Similarity 52.7%; Pred. No. 6.3e-17;

Matches 208; Conservative 0; Mismatches 187; Indels 0; Gaps 0;

QY 606 tctcaaaaagactcgttctgtagaattcaacgagtcacatttcaacatgcgaagcaag 665  
DB 199 TATTAAGAAAGCCCTTTATTTGATGACACAGATGAAACCTTAGGACAACTATAGAAA 258  
QY 666 ggcacccctgtagtgcctgtagtctcaccatccagagctgaggtttatgatacagtgctc 725  
DB 259 AACCGTCTGGCTTGGGCTTTTATAGCCCTATTTATCAAGCTGAAGCTGAGATTAAGTTT 318  
QY 726 tcaatcaacttaagaacatgctcctccatcctgtagtctcctgtagtctgtagtctcctac 785  
DB 319 ATGTACACTTAAAAAACCTTGCCTTAGAGCCCTACACCTTTTACATGGAATTAATT 378  
QY 786 actggaagagctctgagagagctggaatagatgatacagacagcaagcaaggaagaaag 845  
DB 379 ACTATAAGACATATGAGGAGGAGCCATACCTGATTAACACACAGATTTTCAAGAGCAG 438  
QY 846 atgataaagctccctgggaggaagcacaatcatatgctgagagagctcgtgaagaaagaaatg 905  
DB 439 ATGACAAAGTATTTCCAGAGAGACATATACATACATGTTGCTTGCACCTGAAGAACAA 498  
QY 906 gtccaatgagctctgagagagctgtagtctcctcattcattctctcattgtagtctgtagt 965  
DB 499 GTCCGTGGGAGAGAGATGGCAATTGTGTACATGAGATTATACCATTCACCATTTGATGCTC 558  
QY 966 taaaagactgtaattcaagpctcatcgga 1000  
DB 559 CAAAAGTATTTGCTTCAGGATTCATGACCTTTA 593

## RESULT 13

AV653260 634 bp mRNA EST 07-SEP-2000  
LOCUS AV653260



DEFINITION	AV653260 GLC Homo sapiens cDNA clone GICDDJ09 3', mRNA sequence.
ACCESSION	AV653260
VERSION	AV653260.1
KEYWORDS	EST.
SOURCE	Homo sapiens human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	Qian,B., Wu,T., Huang,Q., Huang,C., Kang,B., Gao,X., Xu,Z., Xiao,H., Xu,X., Li,N., Peng,Y., Liu,F., Qu,J., Song,H., Cheng,Z., Qu,J., Zeng,L., Xu,S., Guo,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Yang,Y., Gu,Y., Chen,Z. and Han,Z.
TITLE	Homo sapiens cDNA clone
JOURNAL	Unpublished (2000)
COMMENT	Contact: Zeguang Han Chinese National Human Genome Center at Shanghai 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China Tel: 86-21-50801919(ex 45) Fax: 86-21-50801922 Email: hanzg@chgc.sh.cn This clone is available at CHGC in Shanghai.
FEATURES	Location/Qualifiers
SOURCE	1..634 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="GICDDJ09" /clone_lib="GIC" /rissue_type="corresponding non cancerous liver tissue" /dev_stage="Adult" /lab_host="SOLR" /note="Vector: pbluescript sk(-); Site_1: EcoRI; Site_2: XhoI"
BASE COUNT	213 a 119 c 128 g 172 t 2 others
ORIGIN	
Query Match	9.4%; Score 94.2; DB 10; Length 634;
Best Local Similarity	52.4%; Pred. NO.1.6e-16;
Matches 207; Conservative	0; Mismatches 188; Indels 0; Gaps 0;
OY	606 tgtcacaaaagactctgttgcataagcaatgcacatgcgtgaagccaa 665
Db	203 TATATAGAGAGGCCCTTATCTTCAGTACACAGATGAACAACCTTTAGACACATATAGAAA 262
OY	666 gaccaccctgtagtggtctgctagagtcctcacatccaaagctgaagttttagataagtcg 725
Db	263 AACCGCTGTGGCTTGGGCTTTTGAAGCCCATATATCAAGCTCAAACTGAGATTAAGTTT 322
OY	726 tcatatcacattaaagacatgcttcccatcctgtcaagttcatalgctgttgtagtacct 785
Db	323 ATGTACACTTAATAAAAAACCTTGCCCTTAGAGCCCTACACACCTTATTCACATGGAATTAACCT 382
OY	786 actggaagaagctctgaaggaagctgaatgaatgaatgacagacagtcgaagaaggaagaag 845
Db	383 ACTATATAGAGACATATAGAGGGGCCCATCTACCCCTGATATACCCACAGATTTTCAAGAGCAG 442
OY	846 atgataaagctctccctgtagtggaagcatalatgctctggaaggtcctgaagaagaatg 905
Db	443 ATGACAAAGTATATTCAGAGAGAGCAGTATACATATGTTGCTTGCACATGAAGAACACA 502
OY	906 gtccaaatgacctgacccacactgtgctctactactatcatatcttctctatgtggaacctgg 965
Db	503 GTCTCGGGGAAGAGAAATGGCAATTTGTGTACAGATTTTACCATTCCACATTTGATGCTC 562
OY	966 taaagaacttgaatlcacagccatcattgagcccta 1000
Db	563 CAAAGATATTTCCTCAGAGACTCATTCGAGACCTTTA 597

LOCUS	AA461838	477 bp	mRNA	EST	10-JUN-1997
DEFINITION	V195J04.r1 Soares_mammary_gland.NbMxg Mus musculus cDNA clone IMAGE:851527 5' similar to gb:M13699 CERULOPLASMIN PRECURSOR (HUMAN)) ; mRNA sequence.				
ACCESSION	AA461838				
VERSION	AA461838.1	GI:2186729			
KEYWORDS	EST.				
SOURCE	house mouse.				
ORGANISM	Mus musculus.				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathia; Muridae; Murine; Mus.				
AUTHORS	1 (bases 1 to 477) Marras, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.				
TITLE	The Mashu-HHMI Mouse EST Project				
JOURNAL	Unpublished (1996)				
COMMENT	Contact: Marra M/Mouse Est Project Mashu-HHMI Mouse Est Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: mouseest@watson.wustl.edu This clone is available royalty-free through LLNL; contact the IMAGE Consortium (infoimage.llnl.gov) for further information. MG:503679 Seq primer: -28ml3 rev2 ET from Amersham High quality sequence stop: 413. Location/Qualifiers 1..477 /organism="Mus musculus" /strain="C57BL/6J" /db_xref="taxon:10090" /clone="IMAGE:851527" /clone_1db="Soares_mammary_gland_NbMxg" /sex="male" /tissue_type="mammary_gland" /dev_stage="4 weeks" /lab_host="DH10B" /note="Organ: mammary gland. Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker. Site_1: Not I. Site_2: Eco RI. 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTACCAATCTGTAAGTGGAGCGCGCGAATGTTTTTTTTTTTTTTTTTTTTTTTTTTT T 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia). digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. RNA provided by Dr. Minoru Ko, Wayne State Univ. Library constructed and normalized by Bento Soares and M.Palima Bonaldo."				
BASE COUNT	148 a	102 c	102 g	125 t	
ORIGIN					
Query Match	9.4%	Score 93.8;	DB 10;	Length 477;	
Best Local Similarity	54.5%;	Pred. No. 1.9e-16;			
Matches 210;	Conservative 0;	Mismatches 172;	Indels 3;	Gaps 1;	
608	tacaaaagacgtcttctgtgaatccacgcatcccttcaacaatcgctaaagccaag	667			
Db	77 TATTAAGAAAGCCCTTATTTATTTAGATACACAGTGCACCTTTAGTAAGACTTATAGACAAA	136			
668	ccacccttgatgagctctgtagtccatccacaccacccagctgaggtattatgatcagatgctc	727			
Db	137 CCAGCGTGGCTAGGGGTTTATAGGCCCTGTCAACAAAGCTGAAGTGAAGATAAAGTTAT	196			
678	attaaactaagaagaatgcttccatccatctctcagctctcaatgctctgtgtggtgatctccac	787			
Db	197 GTTCACTTAAAGAACCTTGCCTTAGATCTTACACTTTCATGACATGCGGGGTACGTTAC	256			

OY	788	tgaaagctctcgaaggaactgaataatgltgctcagacagcacaagggagaagaagat	847
Db	257	ACCAAGGAGTATGAGGGAGCCGCTTACCTTGACACACCCACTGATTTTAAAGGGCTGAT	316
OY	848	gataaagctctccctcgtgtgaaagcacaatactgtctgcaagctccgaagaagaatgtg	907
Db	317	GACCAATGTCCTTCCGGACACACGATGCTGATGCTGCATGCC--AATGGCCCACT	373
OY	908	ccaatggcctctgagcccaactgtgccttaactactcatactctctctcatctgtgaactgtga	967
Db	374	CTTGAGAGGAGGAGACGCAATGTGTGACACAGATTTACCACTCCCATGTTATGCTCCA	433
OY	968	aaagactgaaactgaagcctcatctg	992
Db	434	AAAGATATTGCATCAGACTCATAG	458

RESULT	15
BE914540	
LOCUS	BE914540
DEFINITION	musculus CDNA clone IMAGE:3965721 5'
	mRNA sequence.

VERSION	BE914540.1	GI:10413270
KEYWORDS	EST.	
SOURCE	house mouse.	
ORGANISM	Mus musculus	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 576)	
AUTHORS	NIH-MGC <a href="http://mgc.nhl.nih.gov/">http://mgc.nhl.nih.gov/</a> .	
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)	
JOURNAL	Unpublished (1999)	
COMMENT	Contact: Robert Strausberg, Ph.D.	

Tissue Procurement: Gilbert Smith, Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MCC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Plate: IMAGE9137 row: 1 column: 10  
High quality sequence stop: 576.

FEATURES	SOURCE	Location/Qualifiers
		1. 576
		/organism="Mus musculus"
		/strain="FVB/N"
		/db_xref="taxon:10090"
		/clone="IMAGE:3965721"
		/clone_id="NCI-CCAP_Mam1"
		/tissue_type="tumor, biopsy sample"
		/dev_stage="3 months, virgin"
		/lab_host="DH10B"
		/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NciI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigator providing samples: Gilbert Smith, NIH"
BASE COUNT		178 a 119 c 127 g 151 t 1 others
ORIGIN		

Query	Match	Similarity	Score	DB	Length
Best	Local	53.98	Pred. No. 7.9e-16		576
Matches	212	Conservative	0	Mismatches	177
				Indels	4
				Gaps	1
Oy	608	tacaaaagacgcctgtgtgaagatcaccgatacccttcaacatcgtacgaag	667		
Db	157	TATTAAGAAAGGCCCTTATTTTATTTAGTACACAGATGACACCTTTAGTATAGACTATAGACAA	216		
Oy	668	ccaccctgatgagctctcctagtcctccaccacccagcgtcgaagttatatacagtcgc	727		
Db	217	CCAGCTCTGGCTTAGGGTTTATTAGGCCCTGTCATCAAAAGCTGAAGATGGACATTAAGTTAT	276		

Oy	728	atcacactaaagaatgagctcccatccctgtaagctctatagctgttggtatccac	787
Db	277	gTTCACTTAAAGACCTTGCCCTCAGGATCTACACTTTTCATGACATGGGGTAACTAC	336
Oy	788	tgysaaagctctcgtagysagctgaaatgatatgatacagaaccagtcasaaagsaaagaat	847
Db	337	ACCAAGGGATGTAGGGGAGCGGTACCCCTGCACACACCATGTATTTTCAACGGCGGTAT	396
Oy	848	gataaagctctccctcgtgtagaagcatatacatgctgcgcaggtccctgaaagaatggt	907
Db	397	GACAAAGTGTCTCCCGGACACACTATGTGTATGTGCTGCATGCGCAATGACCAAGTCTT	456
Oy	908	ccaatgagctctgacccaagctgtgccttaacctactacatactctctcaatgtagaactgta	967
Db	457	GGAGAGG----GAGACACCAATGTGTGACCCAGATTTACACACTCCATGTTGATGTCACA	512
Oy	968	aaagactgaaatccagagctcatgtagaccca	1000
Db	513	AAAGTATTGCATCAGACGACTATAGGACCTCTA	545

Search completed: January 8, 2002, 15:13:44  
Job time: 4820 sec

Search completed: January 8, 2002, 15:13:44  
Job time: 4820 sec

Wed Jan 9 14:38:29 2002

us-09-740-211-14\_copy\_4000\_4999.rge

Page 1

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 8, 2002, 17:25:06 ; Search time 7258.17 Seconds  
(without alignments)  
2272.912 Million cell updates/sec

Title: US-09-740-211-14\_COPY\_4000\_4999

Perfect score: 1000  
Sequence: 1 gatacagtgatcgtgtgg.....cgcagctgcctgcagacacat 1000

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmb1: \*  
1: gb\_da: \*  
2: gb\_hlg: \*  
3: gb\_in: \*  
4: gb\_om: \*  
5: gb\_ov: \*  
6: gb\_pat: \*  
7: gb\_ph: \*  
8: gb\_pl: \*  
9: gb\_pr: \*  
10: gb\_ro: \*  
11: gb\_sts: \*  
12: gb\_sy: \*  
13: gb\_un: \*  
14: gb\_vl: \*  
15: em\_da: \*  
16: em\_fun: \*  
17: em\_hum: \*  
18: em\_in: \*  
19: em\_om: \*  
20: em\_or: \*  
21: em\_ov: \*  
22: em\_pat: \*  
23: em\_ph: \*  
24: em\_pl: \*  
25: em\_ro: \*  
26: em\_sts: \*  
27: em\_sy: \*  
28: em\_un: \*  
29: em\_vl: \*  
30: em\_higo\_hum: \*  
31: em\_higo\_inv: \*  
32: em\_higo\_rod: \*  
33: em\_hlg\_hum: \*  
34: em\_hlg\_rod: \*  
35: em\_hlg\_inv: \*  
36: em\_hlg\_other: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1000	100.0	4999	6	AR138378
2	1000	100.0	4999	6	AR146888
3	954.4	95.4	11933	6	AR138377
4	954.4	95.4	11933	6	AR146887
5	790	79.0	4670	6	AR110040
6	790	79.0	5035	6	AR034084
7	790	79.0	5035	6	AR071306
8	790	79.0	5035	6	AR071272
9	790	79.0	6999	6	177105
10	790	79.0	7056	6	127063
11	790	79.0	7272	6	105404
12	790	79.0	8667	6	AR052730
13	790	79.0	8967	6	171409
14	790	79.0	8967	6	HSEFVITR
15	790	79.0	9009	6	AR003710
16	790	79.0	9009	6	AR029065
17	790	79.0	9009	6	AR126884
18	790	79.0	9009	6	131901
19	790	79.0	9009	6	163424
20	790	79.0	9009	6	HUMFVITR
21	790	79.0	9029	9	HUMFVITR
22	790	79.0	9329	9	AR003585
23	788.4	78.8	8241	6	A05328
24	788.4	78.8	8241	6	A07042
25	787	78.7	4278	6	108644
26	787	78.7	4551	6	108641
27	786.8	78.7	7440	6	108345
28	785.4	78.5	4281	6	108643
29	785.4	78.5	4548	6	108642
30	783.6	78.4	7440	6	108457
31	695.4	69.5	1728	6	102054
32	634.8	63.5	7493	6	AR003712
33	634.8	63.5	7493	6	AR029067
34	634.8	63.5	7493	6	AR126886
35	634.8	63.5	7493	6	163427
36	634.8	63.5	7493	10	MUSCVCVITR
37	628.4	62.8	2493	3	HUMF8C
38	619.8	62.0	7145	4	AF049489
39	614	61.4	7032	4	AF016234
40	594.8	59.5	4334	6	AR029098
41	594.8	59.5	4334	6	AR126917
42	594.8	59.5	6402	6	AR029097
43	594.8	59.5	6402	6	AR126916
44	594.8	59.5	6539	4	SSU49517
45	592	59.2	8831	6	E00527

#### ALIGNMENTS

RESULT 1  
LOCUS AR138378 4999 bp DNA PAT 16-JUN-2001  
DEFINITION Sequence 14 from patent US 6200560.  
ACCESSION AR138378  
VERSION AR138378.1 GI:14480723  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 4999)  
AUTHORS Couto, L.B., Colosi, P.C. and Qian, X.  
TITLE Adeno-associated virus vectors for expression of factor VIII by target cells  
JOURNAL Patent: US 6200560-A 14 13-MAR-2001;  
FEATURES  
source location/Qualifiers  
BASE COUNT 1375 a 1151 c 1155 g 1318 t  
ORIGIN

Query Match 100.0%; Score 1000; DB 6; Length 4999;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-299;  
 Matches 1000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gatcaagtgatctgttggcacaatgatattcaagcgaatcaagaccagggctgcg 60  
 Db 4000 GATCAAGTGATCTGTGGCACAATGATATTCAAGCATCAAGCCAGGGTGC CGC 4059

QY 61 tcagaagttcccaagctctacatctctcgttatcatatcatatgatagttatggaa 120  
 Db 4060 TCAGAAGTTCTCCAGCCTCTACATCTCTCGATTATCATATGATATGTTATGGAA 4119

QY 121 gaagtgcagactatctgaggaatccacttgaaacttaatgctctcttggaatgt 180  
 Db 4120 GAAGTGCAGACTTATCGAGGAATTCCTACCTGAGACTTAAATGCTCTTTGGCAATGT 4179

QY 181 ggaattcctgggataaaacaatatctttaaaccctccaattatgtctgatalaccg 240  
 Db 4180 GGATTCTCCTGGGATAAAACAATATTTTAAACCCCTCAATATTGCTGATACATCCG 4239

QY 241 ttggaccacactaatatgatgatcttgcaagcactcttgcaatggatgtatggcgtga 300  
 Db 4240 TTGGACCACTCATATATGATGATTCGACACTCTTCGATGGAGTTGATGGGCTGA 4299

QY 301 tttaaatagttgcagcaatgcacattgggaatgagagtaaagcaataacagatcacagat 360  
 Db 4300 TTTAAATAGTTGCAGCATGCTTGGGAATGAGAGTAAAGCAATATCATGATCAGAT 4359

QY 361 tactgtctacactcttcaatctgtttgcaactgtctcctctcaaaagctgcact 420  
 Db 4360 TACTGCTTACTCTTACTTTACCAATATGTTGCCACTGCTCTCTTCAAAACCTGCAGCT 4419

QY 421 tcaactccaagaggagagtaatgctcctggaacactcagtgtaataacaaagagtgct 480  
 Db 4420 TCACCTCCAAGGAGGAGATGATGCTGAGACCTCAGTGAATATCAAAAGAGTGCT 4479

QY 481 ggaagtggaacttcagaagaacaaatgaaatcagaagatgaactactcaaggaataaac 540  
 Db 4480 GGAAGTGGAACCTTCAGAAACAAATGAATGCAGAGTACTACAGGGAGTAAATTC 4539

QY 541 tctgctaccagcatgatgtgaagagttccatctcctcagaagcagatgagcatca 600  
 Db 4540 TCTGCTTACCAGCATGTATGTGAAGAGTCTCTCATCTCCACAGATCAAGATGGCCATCA 4599

QY 601 gtggactctcttttcagaatgcaaaatgaagtttttcaagggaatcaagactcct 660  
 Db 4600 GTGGACTCTCTTTTCAGAATGGCAAGTAAAGTTTTCAGGGGAATCAAGACTCCTT 4659

QY 661 caaacctgtgtgaactctctgagcccaacgcttactgactgctacccctgaattcaacc 720  
 Db 4660 CAAACCTGTGTGAACCTCTGTAGCCACCGTACTGACTGCTCTCTCGAATTCACCC 4719

QY 721 ccaagattgaggaacagatgctcctgagatgaggttttcggagctcgaaggaacaga 780  
 Db 4720 CCAGATTGGGAGCAACAGATGCTTAGAGATGAGAGTTCGGGCTGCGAGGCAAGGA 4779

QY 781 cctctactgacctgagacttaataaagaattatcttcaatgcaatagtggtgtgtt 840  
 Db 4780 CCTTACTGACTGAGACCTTAATTAAGGAATTTATTTTCAATGCAATAGTGTGTTT 4839

QY 841 ttgtgtgctggcaggaacccctgaatgagagttggcaactccctctgagcgtc 900  
 Db 4840 TTTGTGTGGCGCGCAGGAACCCCTAGTATGAGATTG6CACTCCTCTCTGCGCGCTC 4899

QY 901 gctgctactgaagcctgggagcaacaaagtcgcccgaacgcgcgggctttgcccggag 960  
 Db 4900 GCTGCTACTGAGGCGCGGAGCAACAAAGTGCGCCGAGCGCGGCGCTTTGCCGGGCGG 4959

QY 961 cctcagtgagagagagagagagagctgctcctgcaagacat 1000  
 Db 4960 CCTCAGTGAGAGGAGGAGGAGGCGGCACTGCTGCAAGACAT 4999

RESULT 2  
 ARI46888 ARI46888 4999 bp DNA PAT 08-AUG-2001  
 LOCUS Sequence 14 from patent US 6221349.  
 DEFINITION ARI46888  
 ACCESSION ARI46888  
 VERSION ARI46888.1 GI:15110691  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE Unclassified.  
 1 (bases 1 to 4999)  
 AUTHORS Couto, L.B., Colosi, P.C. and Qian, X.  
 TITLE Adeno-associated vectors for expression of factor VIII by target cells  
 JOURNAL Patent: US 6221349-A 14 24-Apr-2001;  
 FEATURES Location/Qualifiers  
 source 1..4999  
 BASE COUNT 1375 a 1151 c 1155 g 1318 t  
 ORIGIN

Query Match 100.0%; Score 1000; DB 6; Length 4999;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-299;  
 Matches 1000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gatcaagtgatctgttggcacaatgatattcaagcgaatcaagaccagggctgcg 60  
 Db 4000 GATCAAGTGATCTGTGGCACAATGATATTCAAGCATCAAGCCAGGGTGC CGC 4059

QY 61 tcagaagttcccaagctctacatctctcgttatcatatcatatgatagttatggaa 120  
 Db 4060 TCAGAAGTTCTCCAGCCTCTACATCTCTCGATTATCATATGATATGTTATGGAA 4119

QY 121 gaagtgcagactatctgaggaatccacttgaaacttaatgctctcttggaatgt 180  
 Db 4120 GAAGTGCAGACTTATCGAGGAATTCCTACCTGAGACTTAAATGCTCTTTGGCAATGT 4179

QY 181 ggaattcctgggataaaacaatatctttaaaccctccaattatgtctgatalaccg 240  
 Db 4180 GGATTCTCCTGGGATAAAACAATATTTTAAACCCCTCAATATTGCTGATACATCCG 4239

QY 241 ttggaccacactaatatgatgatcttgcaagcactcttgcaatggatgtatggcgtga 300  
 Db 4240 TTGGACCACTCATATATGATGATTCGACACTCTTCGATGGAGTTGATGGGCTGA 4299

QY 301 tttaaatagttgcagcaatgcacattgggaatgagagtaaagcaataacagatcacagat 360  
 Db 4300 TTTAAATAGTTGCAGCATGCTTGGGAATGAGAGTAAAGCAATATCATGATCAGAT 4359

QY 361 tactgtctacactcttcaatctgtttgcaactgtctcctctcaaaagctgcact 420  
 Db 4360 TACTGCTTACTCTTACTTTACCAATATGTTGCCACTGCTCTCTTCAAAACCTGCAGCT 4419

QY 421 tcaactccaagaggagagtaatgctcctgagatgaggttttcaagggaatcaagactcct 480  
 Db 4420 TCACCTCCAAGGAGGAGATGATGCTTAGAGACTCAGAGTAAATATCAAAAGAGTGCT 4479

QY 481 ggaagtggaacttcagaagaacaaatgaaatcagaagatgaactactcaaggagtaaac 540  
 Db 4480 GGAAGTGGAACCTTCAGAAACAAATGAATGCAGAGTAAATGCTCAGGAGTAAATTC 4539

QY 541 tctgctaccagcatgatgtgaagagttccatctcctcagaagcagatgagcatca 600  
 Db 4540 TCTGCTTACCAGCATGTATGTGAAGAGTCTCTCATCTCCAGAGTCAAGATGGCCATCA 4599

QY 601 gtggactctcttttcagaatgcaaaatgaagtttttcaagggaatcaagactcct 660  
 Db 4600 GTGGACTCTCTTTTCAGAATGGCAAGTAAAGTTTTCAGGGGAATCAAGACTCCTT 4659

QY 661 caaacctgtgtgaactctctgagcccaacgcttactgactgctacccctgaattcaacc 720

```

Db 4660 CACACCTGTGTAACCTCTGTAACCCACCGTTACTGACTGCTACCTCGAATTACCC 4719
OY 721 ccgaattgggagaccacattggccttaggattgaggttcggagctgcgaagcaaga 780
Db 4720 CCAGAGTTGGGGTACACAGATTGGCCTTAGAGATGAGGTTCTGGGCTGGAGGACAGA 4779
OY 781 cccctactgactcgaagcctaataaagaaattatlttaattgacaaatagtggtgtt 840
Db 4780 CCTCTACTGACTCGAGCCTTAATAAGAAATTTATTTTCATTGCAATAGTGTGTGTTT 4839
OY 841 ttgtgtgagcgagcaagaacctgtatgtatgttgcaactccctctctgagcgctc 900
Db 4840 TTTGTGTGGCGCGCAGAGAACCCCTAGTAGATGAGTTGGCCACTCCCTCTGTGGCGCTC 4899
OY 901 gctcgactcactgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcg 960
Db 4900 GCTGCTGCTGACTGAGCGCGGCGGCGACCAAGTCCGCCGAGCGCGCGGCTTTGCCCGGGCG 4959
OY 961 cctcagtgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcg 1000
Db 4960 CCTCAGTGAGCGAGCGAGCGCGCGAGCTGCTCGAGGACAT 4999

RESULT 3
ARI38377 11933 bp DNA PAT 16-JUN-2001
LOCUS Sequence 13 from patent US 6200560.
DEFINITION ARI38377
ACCESSION ARI38377
VERSION ARI38377.1 GI:14480722
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 11933)
AUTHORS Couto, L.B., Colosi, P.C. and Qian, X.
TITLE Adeno-associated virus vectors for expression of factor VIII by
        target cells
JOURNAL Patent: US 6200560-A 13 13-MAR-2001;
        Location/Qualifiers
FEATURES
        source 1..11933
        /organism="unknown"
BASE COUNT 3258 a 2818 c 2717 g 3140 t
ORIGIN
Query Match 95.4%; Score 954.4; DB 6; Length 11933;
Best Local Similarity 97.9%; Pred. No. 2,2e-285;
Matches 979; Conservative 0; Mismatches 16; Indels 5; Gaps 1;

OY 1 gatcaagtgatctgttggcaccatgattatcaaggcactcaagcccgaggtgccc 60
Db 3975 GATCAAGGTGATCTGTGGACCAATGATTTACCGCATCAAGCCAGGCGTGGCG 4034
OY 61 tcaagaattcctcagcctcactacatctcagttatcatcatatgatatgttgaaggaa 120
Db 4035 TCAGAGTTTCCACACCTCTACATCTCTCAGTTATCATGATGATGATGATGAGAA 4094
OY 121 gaagtggcagactatcgaggaatcactcaggaacttaatgtctcttcttgacatgt 180
Db 4095 GAAGTGGCAGACTTATTCAGAGAAATTCACGTGAGACCTTAATGCTCTTGTGGCAATGT 4154
OY 181 ggaattcatctggataaacaacaataatttaacctccaattatgtctgatacatccg 240
Db 4155 GGATTCATCTGGGATTAACCAATATTTTAAACCTCCATTAATGCTCGATCATCCG 4214
OY 241 ttggacccaactcatlaagcatcgcagcaactctcgatgagatgaggtgtgtga 300
Db 4215 TTTGCAACCAACTATTATAGCATTCGACGACTCTTGCAAGGAGTTGATGGGCTGGA 4274
OY 301 tttaaatgttcacatgacatcttggaattgagagtaagaacatacagatgacagat 360
Db 4275 TTTAAATAGTTCCAGATGCCATTGGGAATGAGAGTAAAGCAATATCATGATGCACAGAT 4334

```

```

OY 361 tactgtcactcactacttaccataatgttggcaccctggtctccttcaaaagctgact 420
Db 4335 TACTGTCTCATCTTACCAATATGTTGGCAACCTGGTCTCTCTTCAAAAGCTCAGCT 4334
OY 421 tcaactcacaaggagaaatagtccttgagacactcagtgtaataatccaaagagtgct 480
Db 4395 TCACCTCCAAAGGAGGAGATTAATGCTGAGACCTCAGGTGAATATCCAAAGAGTGGCT 4454
OY 481 gcaagtgacttccagaagacaatgaaagtcacagagtaactactcagagagtaaatc 540
Db 4455 GCAAGTGACTTCCAGAGACAAATGAAGTCAAGAGTAATCTCAAGGAGTAATATC 4514
OY 541 tctgtttccagcaatgatagttgaaagattcctaattcccaagcagtgcaaga 600
Db 4515 TCTGCTTACCAAGCATGATATGTAAGAGATTCTCTCATCTCCAGACGATCAAGTGGCATCA 4574
OY 601 gtgactctcttcttcagaatgcaagtaaaagttttcaaggaaatcaagactcct 660
Db 4575 GTGACTCTCTTTTTCAGAAATGCAAGTAAAGGTTTTCAGGGAAATCAAGACTCTCTT 4634
OY 661 caacctgtgtgaaactctctagaccacgcttactgactcactccttgaaattcacc 720
Db 4635 CACACCTGTGTAACCTCTTAGACCCACCGTTACTGACTGCTTCAATTTACACC 4694
OY 721 ccagaattggtgccaacgaattgcccctgagatggaagttctgagctgcgaagcaaga 780
Db 4695 CCAGAGTTGGGTGCACACCAATTTGCCCTGAGATGAGAGGTTCTGGGCTCGAGGACAGA 4754
OY 781 cctctactgactcgaagcctaataaagaaattatlttaattgcaatgtgtgtgtt 840
Db 4755 CCTCTACTGACTCGAAGAAATGAAGATCAGAGCTCTAGAGAT----CTGTGTGTGTGTTT 4809
OY 841 ttgtgtgagcgagcgaggaacccctagatgaggttgccactccctctctgcgcgtc 900
Db 4810 TTTGTGTGGCGCGCAGAGAACCTTAGTATGAGATTGGCCACTCCTCTCTGCGCGCTC 4869
OY 901 gctcgtcactgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcg 960
Db 4870 GCTGCTGCTGACTGAGCGCGGCGGCGACCAAGGTCGCCGAGCGCGCGGCTTTGCCCGGGCG 4929
OY 961 cctcagtgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcg 1000
Db 4930 CCTCAGTGAGCGAGCGAGCGCGAGCTGCTCGAGGACAT 4969

RESULT 4
ARI46887 11933 bp DNA PAT 08-AUG-2001
LOCUS Sequence 13 from patent US 6221349.
DEFINITION ARI46887
ACCESSION ARI46887
VERSION ARI46887.1 GI:15110690
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 11933)
AUTHORS Couto, L.B., Colosi, P.C. and Qian, X.
TITLE Adeno-associated virus vectors for expression of factor VIII by target
        cells
JOURNAL Patent: US 6221349-A 13 24-APR-2001;
        Location/Qualifiers
FEATURES
        source 1..11933
        /organism="unknown"
BASE COUNT 3258 a 2818 c 2717 g 3140 t
ORIGIN
Query Match 95.4%; Score 954.4; DB 6; Length 11933;
Best Local Similarity 97.9%; Pred. No. 2,2e-285;
Matches 979; Conservative 0; Mismatches 16; Indels 5; Gaps 1;

OY 1 gatcaagtgatctgttggcaccatgattatcaaggcactcaagcccgaggtgccc 60

```



```
Db 4433 CCTCTACTGAGGTGGCC 4450
|||||
RESULT 6
AR034084 5035 bp DNA PAT 29-SEP-1999
LOCUS AR034084
DEFINITION Sequence 1 from Patent US 5869292.
ACCESSION AR034084
VERSION AR034084.1 GI:58496689
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 5035)
AUTHORS Voorberg,J.J.
TITLE Hybrid proteins with modified activity
JOURNAL Patent: US 5869292-A 1 09-FEB-1999;
FEATURES
Location/Qualifiers
source 1..5035
/organism="unknown"
BASE COUNT 1484 a 1127 c 1110 g 1314 t
ORIGIN
Query Match 79.0%; Score 790; DB 6; Length 5035;
Best Local Similarity 99.4%; Pred. No. 2.7e-234;
Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 gatcaagtgatgctgttgagcacaatgattatcagcgcatcaagaagcaggtgccg 60
Db 4231 GATCAAGTGTGATCTGTGGACCAATGATTATTCACGGCATCAACCCAGGTCGCCG 4290
QY 61 tcagaagttctcagcctcacaatctcagttatcatcatgatagtatcttgatggaa 120
Db 4291 TCAGAAGTCTCAGCCTTACATCTCTCAGTTTATCATCATGATATAGTTCGATGGAA 4350
QY 121 gaagtgacagactatcaggaatccatcaggaacttaagtctctcttggaatgt 180
Db 4351 GAAGTGCACACTTATTCGAGAAATTCACCTGACACCTTATGCTCTCTTTGGCAATGT 4410
QY 181 ggattacatctgggataaacaacataatcttaacccctcaattatgtctcgatata 240
Db 4411 GGATTACTCTGGGATAAACAATATTTTAAACCTCCATATTATTCGATATCCG 4470
QY 241 ttgaccccaactatataagcattcgacacactcttcgatagtgtgagtgctgta 300
Db 4471 TTTCACCCCAACTATATAGCATTCGACACACTCTTCGATGAGTGGGCTGTGA 4530
QY 301 tttaaatagttgacagcatgcatgggaatcgagaglaaagcaatcatcagatgcacag 360
Db 4531 TTTTAATATGTTGCAGCATCCATTGGGAATGAGAGTAAGCAATATCAGATGCACAGAT 4590
QY 361 tactgtctcatctacttaaccaatagtgttgcacactgctctcctcaaaagtcgact 420
Db 4591 TACTGCTTCACTACTTATACCAATATGTTGGCACCTGTCTCTCTTCAAAAGCTGCAT 4650
QY 421 tcacctccaagggagagtaatgcttggagacctcaagtgtaataatcaaaagtggt 480
Db 4651 TCACCTCCAAAGGAGAGTAATGCTTGAGACTTCAGAGTGAATATATCCAAAGAGTGGCT 4710
QY 481 gcaagtgacatccagaagaacatgaagatcagagagtaactactcaaggaagtaaaatc 540
Db 4711 GCAAGTGCACCTTCCAGAAAGAAATGAAGATCAGAGATTAATCAAGTGGCAATCA 4770
QY 541 tctgtctacagcatgtatgtgaagagttctcatctccagcagtgtaagatggacata 600
Db 4771 TCTGTCTTACAGCATGTATGTGAAGAGTTCCTCATCTCCAGCATCAAGTGGCAATCA 4830
QY 601 gtgacactcttttccagaatggcaaglaaagttttcagggaatacaagactcctt 660
Db 4831 GTGACTCTCTTTTTCAGATGGCAAGTAAGTTTTCAGGGAATCAAGACTCTCTT 4890
```

```
QY 661 cacacctgtgtgactctctagaccacccgcttactactgcgtactcgaatcaacc 720
Db 4891 CACACCTGTGTGAATCTCTAGACCCAGCTTACTACTGCTACTTGGAAFTTACCC 4950
QY 721 ccagaattgggtgacacagatctcctggaagatgaggttcttggtcgaggaacaga 780
Db 4951 CCAGATTTGGGTGACACAGATTTGCCCTGAGATGAGAGTCTGGGCTGCAGACAGGA 5010
QY 781 cctctactgactcgagcc 798
Db 5011 CCTCTACTGAGGTGGCC 5028
RESULT 7
AR071306 5035 bp DNA PAT 18-FEB-2000
LOCUS AR071306
DEFINITION Sequence 1 from patent US 5910481.
ACCESSION AR071306
VERSION AR071306.1 GI:7222194
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 5035)
AUTHORS Voorberg,J.J.
TITLE Hybrid proteins with modified activity
JOURNAL Patent: US 5910481-A 1 08-JUN-1999;
FEATURES
Location/Qualifiers
source 1..5035
/organism="unknown"
BASE COUNT 1484 a 1127 c 1110 g 1314 t
ORIGIN
Query Match 79.0%; Score 790; DB 6; Length 5035;
Best Local Similarity 99.4%; Pred. No. 2.7e-234;
Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 gatcaagtgatgctgttgagcacaatgattatcagcgcatcaagaagcaggtgccg 60
Db 4231 GATCAAGTGTGATCTGTGGACCAATGATTATTCACGGCATCAACCCAGGTCGCCG 4290
QY 61 tcagaagttctcagcctcacaatctcagttatcatcatgatagtatcttgatggaa 120
Db 4291 TCAGAAGTCTCAGCCTTACATCTCTCAGTTTATCATCATGATATAGTTCGATGGAA 4350
QY 121 gaagtgacagactatcaggaatccatcaggaacttaagtctctcttggaatgt 180
Db 4351 GAAGTGCACACTTATTCGAGAAATTCACCTGACACCTTATGCTCTCTTTGGCAATGT 4410
QY 181 ggattacatctgggataaacaacataatcttaacccctcaattatgtctcgatata 240
Db 4411 GGATTACTCTGGGATAAACAATATTTTAAACCTCCATATTATTCGATATCCG 4470
QY 241 ttgaccccaactatataagcattcgacacactcttcgatagtgtgagtgctgta 300
Db 4471 TTTCACCCCAACTATATAGCATTCGACACACTCTTCGATGAGTGGGCTGTGA 4530
QY 301 tttaaatagttgacagcatgcatgggaatcgagaglaaagcaatcatcagatgcacag 360
Db 4531 TTTTAATATGTTGCAGCATCCATTGGGAATGAGAGTAAGCAATATCAGATGCACAGAT 4590
QY 361 tactgtctcatctacttaaccaatagtgttgcacactgctctcctcaaaagtcgact 420
Db 4591 TACTGCTTCACTACTTATACCAATATGTTGGCACCTGTCTCTCTTCAAAAGCTGCAT 4650
QY 421 tcacctccaagggagagtaatgcttggagacctcaagtgtaataatcaaaagtggt 480
Db 4651 TCACCTCCAAAGGAGAGTAATGCTTGAGACTTCAGAGTGAATATCAAAAGAGTGGCT 4710
QY 481 gcaagtgacatccagaagaacatgaagatcagagagtaactactcaaggaagtaaaatc 540
Db 4711 GCAAGTGCACCTTCCAGAAAGAAATGAAGATCAGAGATTAATCAAGTGGCAATCA 4770
```

OY 541 tctgtaccagaatgtaatgaagbgtccatccatccagcaatgaatgacatca 600  
 |||||  
 Db 4771 TCTGCTTACAGAGATGATGAGAGGTTCCATCTCCAGCAGTCAAGATGCGCATCA 4830  
 |||||  
 OY 601 gtgagctctctttttccagaatggaagttttccaggaatacgaactcctt 660  
 |||||  
 Db 4831 GTGGACTCTCTTTTTCAGATGCAAGTAAGTTTTCAGGAAATCAAGACTCTT 4890  
 |||||  
 OY 661 caaacctgtgtgaactctctagaccaccgttactgactgcttactcgaattaccc 720  
 |||||  
 Db 4891 CACACTGTGTGAATCTCTTACAGCCACCGTTACTGACTGCTTACTCTTCAATTCACCC 4950  
 |||||  
 OY 721 ccagagttgggtgaccaaatttgccttgagatgagttcttggtctgagagcaaga 780  
 |||||  
 Db 4951 CCAGAGTTGGGTGACCAATTTGCCCTGAGAGTGAAGTGTTCGGGCTGCGAGGACACAGA 5010  
 |||||  
 OY 781 cctctactgactgagcc 798  
 |||||  
 Db 5011 CCTTACTGAGGTGGCC 5028  
 |||||

## RESULT 8

AR112722 5035 bp DNA PAT 16-MAY-2001  
 LOCUS Sequence 1 from patent US 6130203.  
 DEFINITION AR112722  
 ACCESSION AR112722  
 VERSION AR112722.1 GI:14092622  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.

REFERENCE  
 1 (bases 1 to 5035)  
 Vooberg, J.J.

TITLE Hybrid proteins with modified activity

JOURNAL Patent: US 6130203-A 1 10-OCR-2000;

FEATURES  
 Location/Qualifiers  
 1..5035

BASE COUNT 1484 a 1127 c 1110 g 1314 t

ORIGIN

Query Match 79.0%; Score 790; DB 6; Length 5035;

Best Local Similarity 99.48; Pred. No. 2.7e-234;

Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 gatcaagtgatctgttgacccaatgattatcaagcagatcaagaccaggtgccc 60  
 |||||  
 Db 4231 GATCAAGGTGATCTGTTGGCCCAATGATTATTCAGGCGATCAAGACCAGGTGCCCG 4230  
 |||||  
 OY 61 tcaagaagttccagcctcatctctcagttatcaatgattagttgagaa 120  
 |||||  
 Db 4231 TCAGAAAGTTCCAGCCCTCTACATCTTCAGTTATCATGTATGATGAGGAA 4350  
 |||||  
 OY 121 gaagtgcagactatcaaggaatctcaactggaacttaagtctctcttgcaatgt 180  
 |||||  
 Db 4351 GAAGTGCAGACTTATTCAGGAATTCACCTGCAACTTAATGCTTCTTGGCAATGT 4410  
 |||||  
 OY 181 ggaatcatctggatgataaacaacatattttaaccctccaattatgtctgatalaccc 240  
 |||||  
 Db 4411 GGATTCATCTGGATTAACCAATATTTTAACTTCACCTCAATATTTGCTGATCAATCCG 4470  
 |||||  
 OY 241 ttgcacccaactatattatgacatctgagacactctcgatgagttgaggtgta 300  
 |||||  
 Db 4471 TTTCACCCCAACTATTATATGATTCGAGACCTTCGATGAGTGAATGGGCTGTGA 4530  
 |||||  
 OY 301 ttaataagttgcagacatgcatgggaatggaagtaagcaatcaatgacagat 360  
 |||||  
 Db 4531 TTTAAATAGTTGACGATCCATTTGGGAATGAGAGTAAGCAATATTCAGATGACAGAT 4590  
 |||||  
 OY 361 taagtctcatcacttaacaaatgattgtgcaactgttctcctcaaaagctcact 420  
 |||||

Db 4591 TACTGCTTCATCCCTTACCAATATGTTGCCACCTGCTTCTCTTCAAAAGCTCGACT 4650  
 |||||  
 OY 421 tcaactcaaggagaglaatgaccttgagacactcaggtgaataatccaaagatggtc 480  
 |||||  
 Db 4651 TCACCTCCAAAGGAGAGTAATGCTGAGACCTCAGTGAATATCCAAAGAGTGCT 4710  
 |||||  
 OY 481 gaagtgcactctcgaagaacatgaagatgacagagtaactcactcaggaagtaaatc 540  
 |||||  
 Db 4711 GCAGTGAATCTTCCAGAAAGCAATGAAGTCAAGAGTAATCTCAAGGAGTAATATC 4770  
 |||||  
 OY 541 tctgtaccagaatgtaatgaagbgtccatccatccagcaatgaatgacatca 600  
 |||||  
 Db 4771 TCTGCTTACAGAGATGATGAGAGGTTCCATCTCCAGCAGTCAAGATGCGCATCA 4830  
 |||||  
 OY 601 gtgagctctctttttccagaatggaagttttccaggaatacgaactcctt 660  
 |||||  
 Db 4831 GTGGACTCTCTTTTTCAGATGCAAGTAAGTTTTCAGGAAATCAAGACTCTT 4890  
 |||||  
 OY 661 caaacctgtgtgaactctctagaccaccgttactgactgcttactcgaattaccc 720  
 |||||  
 Db 4891 CACACTGTGTGAATCTCTTACAGCCACCGTTACTGACTGCTTACTCTTCAATTCACCC 4950  
 |||||  
 OY 721 ccagagttgggtgaccaaatttgccttgagatgagttcttggtctgagagcaaga 780  
 |||||  
 Db 4951 CCAGAGTTGGGTGACCAATTTGCCCTGAGAGTGAAGTGTTCGGGCTGCGAGGACACAGA 5010  
 |||||  
 OY 781 cctctactgactgagcc 798  
 |||||  
 Db 5011 CCTTACTGAGGTGGCC 5028  
 |||||

## RESULT 9

LOCUS 177105 6999 bp DNA PAT 03-APR-1998  
 DEFINITION Sequence 1 from patent US 5693499.  
 ACCESSION 177105  
 VERSION 177105.1 GI:3013259  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.

REFERENCE  
 1 (bases 1 to 6999)

AUTHORS Yonemura, H., Tajima, Y., Sugawara, K. and Masuda, K.

TITLE Process for preparing human coagulation factor VIII protein complex

JOURNAL Patent: US 5693499-A 1 02-DEC-1997;

FEATURES  
 Location/Qualifiers  
 1..6999

BASE COUNT 2224 a 1503 c 1480 g 1792 t

ORIGIN

Query Match 79.0%; Score 790; DB 6; Length 6999;

Best Local Similarity 100.0%; Pred. No. 2.9e-234;

Matches 790; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 gatcaagtgatctgttgacccaatgattatcaagcagatcaagaccaggtgccc 60  
 |||||  
 Db 6210 GATCAAGGTGATCTGTTGGCCCAATGATTATTCAGGCGATCAAGACCAGGTGCCCG 6269  
 |||||  
 OY 61 tcaagaagttccagcctcatctctcagttatcaatgattagttgagaa 120  
 |||||  
 Db 6270 TCAGAAAGTTCCAGCCCTCTACATCTTCAGTTATCATGTATGATGAGGAA 6329  
 |||||  
 OY 121 gaagtgcagactatcaaggaatctcaactggaacttaagtctctcttggaatgt 180  
 |||||  
 Db 6330 GAAGTGCAGACTTATTCAGGAATTCACCTGCAACTTAATGCTTCTTGGCAATGT 6389  
 |||||  
 OY 181 ggaatcatctggatgataaacaacatattttaaccctccaattatgtctgatalaccc 240  
 |||||  
 Db 6390 GGATTCATCTGGATTAACCAATATTTTAACTTCACCTCAATATTTGCTGATCAATCCG 6449  
 |||||  
 OY 241 ttgcacccaactatattatgacatctgagacactctcgatgagttgaggtgta 300  
 |||||



```

Db 6450 TTTGACCAACTCATTTAGCATTCGACGACCTCTTGCATGAGTTCATGGGCTGTGA 6509
Qy 301 ttttaaatagttcgaagacgcttttgggaatggagatgaacaatatacagaatgacagat 360
Db 6510 TTTAAATAGTTGAGCATGCCATTTGGGAATGGAGTAAAGCAATATCAGATGACAGAT 6569
Qy 361 tactgttcaatccttaacttaacaaatattgtttccaccctgtgtccttcaaaagctgact 420
Db 6570 TACTGCTTACCTACTTATACCAATATGTTTCCACCTGCTCTCTTCAAAAGCTGCAGCT 6629
Qy 421 tcacccccaaggagagatattgctgagaccctcaagtgaaataacaaagagtgct 480
Db 6630 TCACCTTCCAAAGGAGAGATATGCTGAGACCTCAGGTAATATCCAAAAGATGAGCT 6689
Qy 481 gcaagtgagacttcgaagaacaaatgaagaatgaagagatgaactaactcagagagtaaatc 540
Db 6690 GCAAGTGGACTTCGACAAACAAATGAAGTTCACAGAGATTAATCTCAGGAGATGAATC 6749
Qy 541 tctgttaccagatgatgtgaagaggttccctcaatcccaagtcgaatgagcatca 600
Db 6750 TCTGCTTACAGCATGTATGTGAAGAGTTCCTCATCTCCACAGATCAAGATGAGCATCA 6809
Qy 601 gtgagactcttttccagaatggcaagtaaggttttccaggaatccaaagactcctt 660
Db 6810 GTGGACTCTCTTTTTCAGAAATGCAAGATGAAGTTTTCAGGGAATCAAGACTCCTT 6869
Qy 661 cacacccgtgtgaactctctagaaccacacgttactgaactcgaactcgaattaccc 720
Db 6870 CACACCTGTGTGTAACCTTATACCCACCGTTACTGACTGCTTCAATTCGAAATTCACCC 6929
Qy 721 ccagagttggttgacacagattgcccctgagagatgagagttctggagctgcagagacaaga 780
Db 6930 CCAAGTTGGGTGACACAGATGCTCCCTGAGAGATGAGAGTTCGCGGTGAGGACACAGGA 6989
Qy 781 ccttactga 790
Db 6990 CCTCTACTGA 6999

RESULT 10
LOCUS 127063 7056 bp DNA PAT 06-FEB-1997
DEFINITION Sequence 1 from patent US 5563045.
ACCESSION 127063
VERSION 127063.1 GI:1817839
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 7056)
AUTHORS Piltman,D., Rehmentulla,A., Wozney,J.M. and Kaufman,R.J.
TITLE Chimeric procoagulant proteins
JOURNAL Patent: US 5563045-A 1 08-OCY-1996;
FEATURES
SOURCE 1..7056
/organism="unknown"
BASE COUNT 2232 a 1517 c 1491 g 1816 t
ORIGIN
Query Match 79.0%; Score 790; DB 6; Length 7056;
Best Local Similarity 100.0%; Pred. No. 2,9e-234;
Matches 790; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 gatcaagtgagatctgttgcgacaaatgatatcaagcgacatcaagaacccaaggtgccc 60
Db 6267 GATCAAGTGGATGCTGTGGGACCAATGATATTCACGCGATCAAGCCAGGCTGCCCG 6326
Qy 61 tcaagaagttctccagccttaaatcttcagttatacatalatgatagttcttgatggaa 120
Db 6327 TCAGAGATTCTCCAGCTCTACATCTCAGTTATATCATATGATATGTTGATGGAA 6386

```

```

Qy 121 gaagtgcaagacttatcgaggaatctccactggaaccttaatggtctcttggcaatgt 180
Db 6387 GAAGTGGCAGACTTATTCGAGAAATTCACACTGGAACCTTAATGCTCTTGTGGCAATGT 6446
Qy 181 gattcattctggataaacaacaatatctttaaaccctcaattatgtctcgataacacg 240
Db 6447 GATTTCATCTGGGATTAACCAATATTTTAAACCTTCATTTATTCGATACATCCG 6506
Qy 241 ttgtaaccaactcaattatagatctcgaagacactcttcgtatggatgttgaagctgtga 300
Db 6507 TTTGCAACCAACTCATTTATAGCATTTTCGACGACTCTTGCATGGAGTTGATGGCTGTGA 6566
Qy 301 ttttaaatagttcgaagatccattgggaatggagatgaagaatgaatgaatgacagat 360
Db 6567 TTTAAATAGTTGAGCATGCCATTTGGGAATGGAGTAAAGCAATATCAGATGACAGAT 6626
Qy 361 tactgttcaatccttaacttaacaaatattgtttccaccctgtgtccttcaaaagctgact 420
Db 6627 TACTGCTTACCTACTTATACCAATATGTTTCCACCTGCTCTCTTCAAAAGCTGCAGCT 6686
Qy 421 tcacccccaaggagagatattgctgagaccctcaagtgaaataacaaagagtgct 480
Db 6687 TCACCTTCCAAAGGAGAGATATGCTGAGACCTCAGGTAATATCCAAAAGATGAGCT 6746
Qy 481 gcaagtgagacttcgaagaacaaatgaagaatgaagagatgaactaactcagagagtaaatc 540
Db 6747 GCAAGTGGACTTCGACAAACAAATGAAGTTCACAGAGATTAATCTCAGGAGATGAATC 6806
Qy 541 tctgttaccagatgatgtgaagaggttccctcaatcccaagtcgaatgagcatca 600
Db 6807 TCTGCTTACAGCATGTATGTGAAGAGTTCCTCATCTCCACAGATCAAGATGAGCATCA 6866
Qy 601 gtgagactcttttccagaatggcaagtaaggttttccaggaatccaaagactcctt 660
Db 6867 GTGGACTCTCTTTTTCAGAAATGCAAGATGAAGTTTTCAGGGAATCAAGACTCCTT 6926
Qy 661 cacacccgtgtgaactctctagaaccacacgttactgaactcgaactcgaattaccc 720
Db 6927 CACACCTGTGTGTAACCTTATACCCACCGTTACTGACTGCTTCAATTCGAAATTCACCC 6986
Qy 721 ccagagttggttgacacagattgcccctgagagatgagagttctggagctgcagagacaaga 780
Db 6987 CCAAGTTGGGTGACACAGATGCTCCCTGAGAGATGAGAGTTCGCGGTGAGGACACAGGA 7046
Qy 781 ccttactga 790
Db 7047 CCTCTACTGA 7056

RESULT 11
LOCUS 105404 7272 bp PAT 02-DEC-1994
DEFINITION Sequence 3 from Patent EP 0265778.
ACCESSION 105404
VERSION 105404.1 GI:591046
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 7272)
AUTHORS Sarver,N. and Drohan,W.
TITLE Factor VIII-C analogs
JOURNAL Patent: EP 0265778-A1 3 04-MAY-1988;
FEATURES
SOURCE 1..7272
/organism="unknown"
BASE COUNT 2271 a 1591 c 1533 g 1877 t
ORIGIN
Query Match 79.0%; Score 790; DB 6; Length 7272;
Best Local Similarity 99.4%; Pred. No. 3e-234;
Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

QY 1 gataaagtgatctgttggcaccatgatattcaagcgcataaagaggtgccc 60  
 Db 6267 GATCAAGGTGATCTGTGGACCAATGATTAATTCAGGCGATCAAGCCAGGTCGCCG 6326  
 QY 61 tcagaattcccaagcctcctacatctcgaattatcaatcagatgtagcttgaggaa 120  
 Db 6327 TCAGAGTTCTCCAGGCTTACATCTCTGAGTTATCATCATCTATATGCTTGATGGAA 6386  
 QY 121 gaagtggcaactatcagaagaattccacttgacccttaagtgtcttcttggaatgt 180  
 Db 6387 GAAGTGGCAGACTTATGAGGAATTCGACTGGAACCTTAATGCTTTCTTGGCAATGT 6446  
 QY 181 ggaatcctcggagataaacaacatattttaacccctcaatattgtcgtgataccg 240  
 Db 6447 GGAATTCATCTGGGATAAACAATATTTTAACTCCCAATATGTCGATACATCG 6506  
 QY 241 ttccaccaactcaatataagcattctcgaagcactctcgaaggagttgagctgtga 300  
 Db 6507 TTTGCACCCAACTATTATAGCATTTGCGACACTCTTCCATGAGATTGATGGCTGTA 6566  
 QY 301 tttaaatagltgacagatgcatttggaatggagaaatgaacatacatcagatgcagat 360  
 Db 6567 TTTAAATAGTTGACAGATGSCATTGGGAATGGAGACTAAAGCAATATCAGATGCACAT 6626  
 QY 361 tactgtctatcctactttaaccataatgttgcacactggtctccttcaaaagctgact 420  
 Db 6627 TACTGCTTCACTCTACTTACCAATATGTTTCCACCTGGCTCCTTCAAAAGCTGACT 6686  
 QY 421 tcaactcgaaggagagaaatgacccctgagacactcagtgaaataatcacaagaatgct 480  
 Db 6687 TCACCTCCAAAGGAGAGATTAATGCTCGAGACTCAGATGATTAATCCAAAAGACTGCT 6746  
 QY 481 gcaagtgagactcccaagaacaaatgaatgaacagagtaactactcagaggaatgaatc 540  
 Db 6747 GCAAGTGAGACTTCCAGAAACAAATGAATGACAGAGTAATCTCAGGAGTAATAATC 6806  
 QY 541 tctgctaccagaatgataatgaaagagttcctcctccagcagtaagtgagatca 600  
 Db 6807 TCTGCTTACACAGATGATGATGAAGAGTCTCTCTCTCCAGCTGCAAGATGAGCATCA 6866  
 QY 601 gtgagactccttttccagaatggaacaaatgaatgaatgaatgaatgaatgaatgaat 660  
 Db 6867 GTGGACTCTCTTTTTCAGAAATGCAAAATGAAGTTTTCAGGGAATCAAGACTCCTT 6926  
 QY 661 caaacctgtgtgaactcctcagaccacccgttaactgctcgaactcgaatcaacc 720  
 Db 6927 CACACTGTGTGAACCTCTTAGACCCACCGTTACTGCTGCTTCAATTCACACC 6986  
 QY 721 ccagaagtgtgacacacagatgcccctgagagtgagagttcctgctcgaaggaacagga 780  
 Db 6987 CCAGAGTTGGGTGACACAGATTGCCCTGAGATGAGATGAGTTCTGGGCTGCGAGCAGGA 7046  
 QY 781 cctctactgactcgaagc 798  
 Db 7047 CCTTACTGAGGCTGCC 7064  
 RESULT 12  
 LOCUS AX052730 8967 bp DNA PAT 12-JAN-2001  
 DEFINITION Sequence 1 from Patent WO0071714.  
 ACCESSION AX052730  
 VERSION AX052730.1 GI:12226920  
 KEYWORDS  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 REFERENCE 1 (bases 1 to 8967)  
 AUTHORS Saenko,E.L. and Strickland,D.K.  
 TITLE Methods of reducing factor VIII clearance and compositions therefor  
 JOURNAL Patent: WO 0071714-A 1 30-NOV-2000.

The American National Red Cross (US)  
 Location/Qualifiers  
 1. 8967  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 110. 166  
 110. 7165  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAC21811.1"  
 /db\_xref="GI:12226921"  
 /translation="MOETLSVCFCLLRFCFSARRRYICAVELSDMYQSDGLP  
 VDAREPPRPKSEPTNTSVYKKTLFLVFEEDNRSMTLNIAPRPMMGLPTTIOEVD  
 VITTKNMAHSVSHAAVGSVSWASBEAETIDDTSOEKEDDVFGSGSHYVVOVL  
 KENGPAADPLCLTSYLSHVDLVKDLNLSGLALVLCRESGLAKETOTLHKFTILF  
 AVDEGKSWHSEKNSLMODDASARMPFMHNGVYVNSRLGLIGCRKSVYMHV  
 IOMGTPEVHSIFLEHPTFLRNHQALESIPITPTLTDLMDLQDFLFCFHS  
 QHDGMEAVYKVDSCPEEPQLMKNNNEEDYDDDTLQSEMDYVRFDDNSRFLQIS  
 VAKHPKTMVHYIAAEEDMDYAPLVLAADRSYKSOYLNNQPORIGRKRYKPMAY  
 TDEPKTREATQHESGILGPIVGEVDLTLLIFKNASRPNIYPIGLTDVPLYSR  
 RLPKGVKHLKOPILPGETFYKATVYEDGPTSDRCULTRYISVFNNEIRLASGL  
 ICPILCTRESYDQKNOIMSKNNVLLSVFEDNRSMTLNIAPRPMMGLPTTIOEVD  
 PEFQASNTMSHNSINGVEDSLQSLVCLHEVAIYLLISGAOTDLSVFFSGYTRKHMV  
 YEDTLTLPFSGEYTFMSENGPLGAILGCHNSDRNRNGATLAKVSCDKNQDXYE  
 SYEDISAYLLSKNNAIEPRSEONSRRSTROKOFNATTPENDIEKTPDMFAHPTM  
 PKIOWSSDLMILROSPPHGLSLSDLOEAKETPSDDSPGADISNLSSEHFE  
 ROLHSGDMVFTPESGOLRLNKLCTTATLAKRIDFVYSTNNLSTIPSDLA  
 ACTDNTSLGSPSMPIYHDSODTLTLLGKSKSLPTEGSGPLSLEENNSKLLSCGLM  
 NSQESSKKNVSTESGRLRGKRAHBPALITKDONALFKVSLILKTNKNSNATNR  
 KTHIDGPLLLENSPWNQNTLESDTEFKVTPILHDMIMDKNATLRLNHSNKT  
 SSKNMEYQOKKEPIPPDAQNPDMSEFKMLIPESARMIORFHGKNSINGGSPK  
 OLVSIGPEKSEVCONFLSEKKNVYVGEFTKVDGKEMVPPSRMLPTLNIDNLEN  
 NTHOENKIOEIEKKEETLIDQENVLPQIHTVGTKNKMLPLSTRONVBSYDCA  
 YAPVLQDRSLNDSTNRKKTTHAKSKKEGEMLEBIGNOTKIVKTACTRIQISNT  
 SOONFYORSRALAKORLPLETELEKRIIVDTSTOMSKNKHLPSTLQIDNE  
 KEGATIOSPLSLDLTRSHSTPOANRSLPIIAVVSFSPSIRPLYTRVLFDONSSHP  
 AASRKKDSVOESSEHPLGAKNNLSLILTEEMGDREVSLSGTSATNSVYKKV  
 ENTVLPKDELPKTSIGVELPKVILYIKDLPLETNNGSGHLDVEGSLGTGTAI  
 KMNANRPGKVPFLVATESSAKPSKLPDLPLAMDNHGYTOIPIKEMKSOEKSEPTA  
 FKKRDTILSNACESNHAIAINDGCONPREITVAKKOORELCSQNPVLRKORE  
 ITRTLOSDEIDYDDTISVEMKKEPDIYDDENOSPSROKTRHFIYAVERLM  
 DYGNSSSHVLRNRAOGSVPOKRVVFOETDGSSTOPLYREGLNHLGLGPIRA  
 EVDNIINWTEKNOASRPYSFSLISVEEDNOGAPBRNPNKPNTEKTYFVKVOHIM  
 APTDEIDECQKMAFYSDVDLEKVDHSLGIPLVCHTNLNAHGHOVVOVOLFET  
 IFDETKSWFTENNERNCRAPCINOEMDEPFKNYEFHAINGYIMDTLPGLMADOR  
 IRWVLSMGSNENHSHFSGHYFVKKREXYMALXNYPCVFEVEMVPSKATIR  
 VECILIGHLHAGMSTLEIYVSNKCPRLCMASHITDPOITASGOTGOMAPRLAHY  
 SGTNASTKEPESWIKYVDLAPMITHGIRTOGAROKESLVIISOFTIYSLDGKMQ  
 TYRNSGTLMVFEFNDVSSGIKHNFNPPIIARYIRLHPHYSLRSTLRMLMGCDL  
 NSCMPLGMESSKAIISDAQITASSYFTNMEKPSKARHLQGRNANPPOVNNPKEN  
 LQVDFQTKMKTGTTQGVKSLTSSVMEFSLISSODSHQWTLFFQNGKVFQGNQ  
 DSFPPVNSIDPLPLTRYLRIRHPQSWHGIALRMEVLGEADLY"  
 mat\_peptide 2841 a 1898 c 1833 g 2395 t  
 BASE COUNT  
 ORIGIN  
 Query Match 79.0%; Score 790; DB 6; Length 8967;  
 Best Local Similarity 99.4%; Pred. No. 3.1e-234;  
 Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 gataaagtgatctgttggcaccatgatattcaagcgcataaagaggtgccc 60  
 Db 6376 GATCAAGGTGATCTGTGGACCAATGATTAATTCAGGCGATCAAGCCAGGTCGCCG 6435  
 QY 61 tcagaagtcccaagcctcctacatctcgaattatcaatcagatgtagcttgaggaa 120  
 Db 6436 TCAGAGTTCTCCAGGCTTACATCTCTGAGTTATCATCATCTATATGCTTGATGGAA 6495  
 QY 121 gaagtggcaactatcagaagaattccacttgacccttaagtgtcttcttggaatgt 180  
 Db 6496 GAAGTGGCAGACTTATGAGGAATTCGACTGGAACCTTAATGCTTTCTTGGCAATGT 6555



## FEATURES

source

misc\_feature

sig\_peptide

CDS

Location/Qualifiers  
 1..8967  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 <1..109  
 /note="5' untranslated region"  
 110..166  
 /product="signal peptide (aa -19 to -1)"  
 110..7165  
 /note="factor VIII precursor"  
 /codon\_start=1  
 /protein\_id="CA25619.1"  
 /db\_xref="GI:31499"  
 /db\_xref="SWISS-PROT:P00451"  
 /translation="MOELSTCEFLCLFFECFSAATRRYLYGAVELSMYDMSDLGELP  
 VDRPPRPVRSFPTSVYKKTLFEVETDHLFNKARPPMKMLGTTIAYVDT  
 VVTLTKMASHPKSLAVGVYKASGEAYDDDTORREKEDVPGSGSHVYQVLT  
 KENGPRASDPLCLITYSLSHVDLYKDLNSGLIGALLVCEGSLAKETQTLTKLILF  
 AVDEKSMHSETKNSLMODRNASARPKMTVNGVNRSLPLGLGCHKSVYWH  
 IGMGTPPEVHSLFEGLTELVRNRQASLEISPTFLAQTLMDLGOFLKCHTISH  
 OHDMGAYKVDSCPEEPOLRMKNNEADYDDDLDSMDVVRPDDKSPFTGIS  
 VAKHRTVWHYIAAEEEDMDYAPLVLDNRYSKSOYLNNNGRGRKKYKRFAY  
 TDEFTKREALHESGLIGLILYGEVDLTLIFKQASRPNTVYHGTIDRPLRY  
 RLKPKYKHLKDPILPEGLFKRYTYVVDGPKSDPRLTQIFSYFVMEEDLSGL  
 IGPLLCYVESVDQNGQIMSDKRVNLFSVDENSRWLTENTQIFLPNPGVLED  
 PERQASNIHMSINGVYVDSILOSLVLAHVAWYILISGAQDTFLSVFSGYFKHKM  
 YEDTLTLFPFSGETVPMSEBNGMLGCHNSDFRNGMTALIKYSCKDNGDYED  
 SYEDISAYLLSKNNAIEPRFSQNSRHRTRKOFNATTIPENDLEKTPFAHRTPM  
 PKTONYSSDLMLLROSPTPHGLSLDLOFAKYEFDSPDPSGATIDSNLSLPMTHF  
 RQHLHSGMWTPPESGLOLRINEKLTGTTAATLEKLDKRVSTSNLSTIPSDNLA  
 AGTDNTSLGPPSPVHYDSQDITLFGKSSPLVESGGLPSENNNSKLLSEGLM  
 NSQESSWGKNVSTESGRLEKGRAPALPLKDNALFVSLTKNTKSNKSLGML  
 KTHIDPSPKLIENSVMONILSDTEFEKVPPLIDRMIMKRNATALLNMSKRT  
 SKNMEVQOKEGPIPPDAONDMSEFPEFLPEBSARIOTRHKNSLNSGQSPK  
 QLYSLGPEKSVESGONFLSEKNKYVVGKGFETDVGLEKEMVPSNRNLPLTNJAHN  
 NTHNOEKKIOEIEKEKTLIOENVYLPQIHYTGKRNKPNFLSTRNVGSDGA  
 YAPVLDPRSLNDSTNRKTHNHSKKEEENLEGLQTOIYEKTAICTQIRISPT  
 SOQENVYORSKRALQKQFLPLETELEKRIIVDDYSTQMSKMKHLTPSLTQIDYNE  
 KEGALITQSPDCLTRSHSIPOANRSPILAKVSFSPISPIYLRVLFQDNSSHL  
 AASYRKDSGVQESHFLOGAKKNLSLILFLMGDQREGLSGTSATNSVYKVV  
 ENT/LKPLDPTKSGVELLPVHIYOKNLOLPLETMSNGSGHLDVEGSLQGTGAT  
 KMEANRPCKVPLRVATRESAKTPKLEIDPLAMNHYQTOIPKEMSOEBSPEKTA  
 FKKTIDILSNBESNHATAINEGONKEIDETAMNKOQRTRELSONPVLKRORE  
 IRTTLQSDOEIDIDDTISVEKKEDPIDEDENQSRQKQTRHETFAVLRRLW  
 DYGMSSPYLRNRAQSGSVPOFKVVEGDEGQKQKDETFYKGEINHEHGLGGLPYIRA  
 EVEDINMTFRNDQSRPYSFSLISYEDDQKQKDETFYKGEINHEHGLGGLPYIRA  
 APTKDEPCKANAYESDVLDEKDHSGLLGPLVCHTNLNAHROLYVOEFLPET  
 IPDETSMYFTENMERNCAPCNICMEDETEKENRPHAINYINDTLDPLGAMADOR  
 ZKRYLLSGNSNENIHSIRSGHYFYRKKEEYKMLNLYRPFVEVEKLPKAKIIR  
 VECILGEHLHAGSTLPLVYSNKCQPLDMASGHLRDFITASGOYQWAPLKLHY  
 SCSINAMSTRKEPSPWIKVLDLAPMIHGIKTQGAOKESLISQITMYSIDGKRW  
 TYRGNSTGLMWFVGNVDSGKIHNFNPILARYIRLPHYISIRSLRMLKCDL  
 NSCSMPLGMSKASIDQAITASSYFTNMAFTPSKARLHROGNSNAMPPOVNNKCM  
 LOVDPOKTMKVTGVTGQVKSLLTSYVVEFLISSQDSOCHQWTLFPOKNAKAVPQNG  
 DSFTPVYNSLDPLPLRLRYLRHPSQVHDIALRMEYLCEADDTL"

mat\_peptide  
 167..7162  
 /product="mature factor VIII (aa 1-2332)"  
 7163..8967  
 /note="3' untranslated region"  
 8948..8953  
 /note="polyadenylation signal"  
 8967  
 /note="polya site"  
 1898 c 1833 g 2395 t

BASE COUNT 2841 a 1898 c 1833 g 2395 t

ORIGIN

Query Match

Best Local Similarity 99.4%; Score 790; DB 9; Length 8967;

Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 gatcaagtgagatctgtgaccacatgattatccagcatcaaacacagtggtccg 60  
 |||||||

Db 6376 GATCAAGTGGATCTGTGGACCAATGATATTCACGGCATCAAGACCAGGTGCCG 6435  
 OY 61 tcagaagttccagccttacatctcgaagtttcacatgatactctgaaggaa 120  
 Db 6436 TCAGAAGTTCTCAGCCTTACATCTCTCAGTTTATCATCATGATATAGCTTGAAGGAA 6495  
 OY 121 gaagtgacagacttacgagaatccacatgaaaccttaagtcctctcttgacaagt 180  
 Db 6496 GAAGTGCAGACTTATCAGAGAAATTCACCTGAACTTAATAGCTCTTTGGCAAGT 6555  
 OY 181 ggaatcatctggagataaacaacaataattttaacctccaatattctcgatataccg 240  
 Db 6556 GAATTCATCTGGGATAAACAACAATATTTTAAACCTCCAAATATTCCTCGATATATCCG 6615  
 OY 241 ttgacacccaactcatatagaatccgacacctcttcgaaggagttgatgggcttga 300  
 Db 6616 TTTGACCCCAACTCATATTATAGCATTCGACGACTCTTGCGATGAGGTTGATGGCTGTGA 6675  
 OY 301 tttaaatagttgcagcatgcca ttggaa ttggaagtaagaaatatacaga tgcacagat 360  
 Db 6676 TTTAAATAGTTGCAGCATGCCATTTGGGAATGAGAGTAAGCAATATCAGATGCACAGAT 6735  
 OY 361 tactgttcaactcacttaccataatgtttgcaacctgtctccttcaaaagctcgact 420  
 Db 6736 TACTGCTTCACTCTACTTATACCAATATGTTTGCACACTGTCTCTCTTCAAAAGCTCGACT 6795  
 OY 421 tcaacctcaaggagagagtaagtcgcttgagagacctcaagg tgaataatcaaaagatgact 480  
 Db 6796 TCACCTCCAAAGGAGAGTAATGCTCGAGACTTACGATGAATATCCAAAGACTGCT 6855  
 OY 481 gaagtgagacttcgaagaacaa tgaaga tgcacagaglaactactcaggaagtaaaatc 540  
 Db 6856 GCAAGTGACTTCGCAAGAGACAAATGAAGTCAACAGAGTAACACTCAGGAGTAATC 6915  
 OY 541 tctgtctaccacatgatatgtgaagagttctcactcctccagcgatcaagtgtgcacatca 600  
 Db 6916 TCTGCTTACACATGATATGTGAAGAGACTTCTCTCATCTCCACACTCAAGATGCCACATCA 6975  
 OY 601 gtcgactcctcttcttcgaatgycagaag taaagttcttcagggaaatcaagactcct 660  
 Db 6976 GTGAGACTCCTTTTTCAGAAATGCAAGTAAGTTTTCAGGAAATCAAGACTCCTT 7035  
 OY 7036 caaacctgtgtgaaactcctctagaccacacgcttactgactcgactccttgaaatcaccc 720  
 Db 7036 CACACTGTGTGAAACTCTCTAGACCCACCGTTACTGACTGCTACCTTCGAATTCACACC 7095  
 OY 721 ccaagattgggagcagcagatgacctgaagatggaagttcttgagctgcagggacagga 780  
 Db 7096 CCAGATTGGGAGCACCGATTGCCCTGAGATGAGAGGTTCTGGGCTCGAGGACACAGGA 7155  
 OY 781 cctctactgactcagacc 798  
 Db 7156 CCTCTACTGAGAGGTGCC 7173

## RESULT 15

AR003710 AR003710 9009 bp DNA PAT 04-DEC-1998  
 LOCUS  
 DEFINITION Sequence 1 from patent US 5744446.  
 ACCESSION AR003710  
 VERSION AR003710.1 GI:3964969  
 KEYWORDS

## ORGANISM

Unknown.  
 Unkown.  
 Unclassified.  
 REFERENCE 1 (bases 1 to 9009)  
 AUTHORS Lollar,J.S. and Runge,M.S.  
 TITLE Hybrid human/animal factor VIII  
 JOURNAL Patent: US 5744446-A 1 28-Apr-1998;  
 FEATURES Location/Qualifiers  
 source 1..9009

BASE COUNT 2853 a 1907 c 1844 g 2405 t

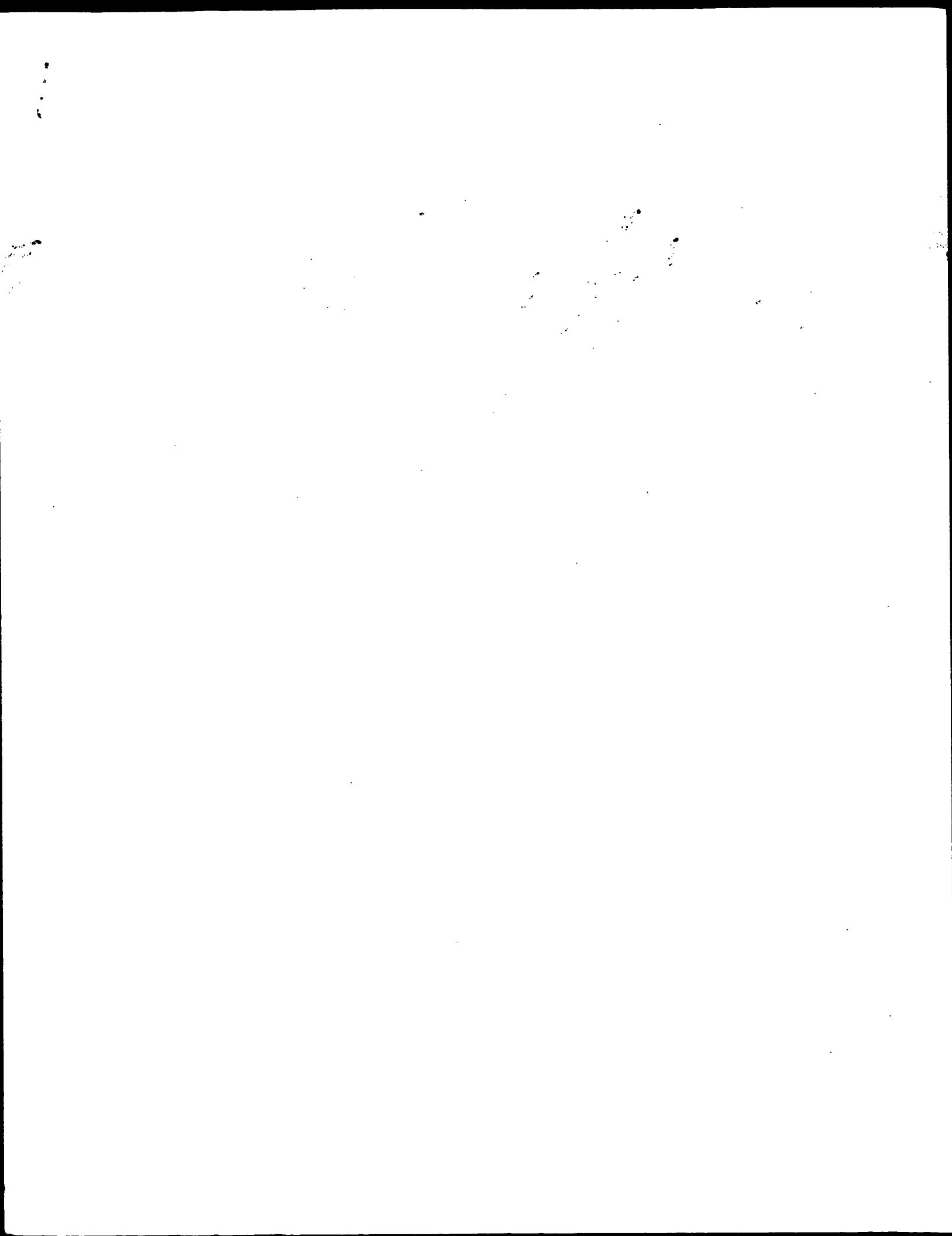
## ORIGIN

Query Match 79.0%; Score 790; DB 6; Length 9009;

Best Local Similarity 99.4%; Pred. No. 3,1e-234; Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
QY 1 gatcaaggtggtatctgttgccacaaatgattatccagcagcaagaccaggtgccc 60
DB 6417 GATCAAGGTGGAATCTGTGTGGCAATGATTTTTCACGCGCATCAAGACCCAGGTGCCG 6476
QY 61 tcagaagttccacagctcaccatctcagttatccatgataagttatgctgagga 120
DB 6477 TCAGAAGTTCTCAGGCTCTACATCTCTCATGTTATCATCATGATAGCTTGATGGGA 6536
QY 121 gaagtgccagacttatccgaggaattccactcagccttaagtctctcttgcaatgt 180
DB 6537 GAAGTGGCACTTATCGAGGAATTCACCTGGAACCTTAATGCTCTCTTGCAATGT 6596
QY 181 gattcatctggagataaacaacaatatcttaaccctccaattatgctgataatccg 240
DB 6597 GATTTCATCTGGAGATAAACAACAATATTTTAACTTCACTTCAATATTTGCTCGATACATCCG 6656
QY 241 ttgcaaccacatcatatagcattcgacacactctcgacatgagttgagtgctgta 300
DB 6657 TTTGCACCAACTCATTTATAGCATTCGACGACTCTTCGATGAGTTGATGGCTGTGA 6716
QY 301 tttaaatagttgcagcagatgcatctggagatgagataagcaataatcagatgcaagat 360
DB 6717 TTTAATAATGTTGCAGATGCCATTTGGGAATGGAGATAAGCAATATCAGATGCACAGAT 6776
QY 361 tactgtctcatctactacttaccataatgtttgcacactgctcctccaagctcgact 420
DB 6777 TACTGCTTATCTCTTACTTACCAATATGTTGCCACCTGCTCTCTTCAAAAGCTCGACT 6836
QY 421 tcaactcgaaggagagaaatgaatgagacatcaggtgaataatccaaagagtgct 480
DB 6837 TCACCTCCAAAGGAGAGATAATGCTCGAGACTCAGGTGATATATCCAAAAGAGTGGCT 6896
QY 481 gcaagtggactccaggaagacaaatgaagtcaagaatlaactactcaggaatgaatc 540
DB 6897 GCAAGTGGACTTCCAGAAAGAAATGAAGATCACAGAGTAACTACTCAAGGAGTAAATC 6956
QY 541 tctgttaccagcatgatagtgaagagttccctcatctccagcaatcaagaatgccaatca 600
DB 6957 TCTGCTTACCAACAGATGATGTAAGAGATTCTCATCTCCAGCACTCAAGATGGCATCA 7016
QY 601 gtggaactctcttttccagaaatgcaagtaaggttttccaggaatacaagactcct 660
DB 7017 GTGGACTCTCTTTTTCAGAAATGCGCAAGTAAAGSTTTTCAGGGAATCAAGACTCCTT 7076
QY 661 caccactgttgatgaactcttgaagccagcttactgactcgtactccttgaaatcaacc 720
DB 7077 CACACCTGTGTGAATCTCTTAGACCACTGACTGACTGACTGACTGACTGACTGACTGACT 7136
QY 721 ccaagattggtgacacagatgctcctgaagatgaggtttcgtgctcgagacaga 780
DB 7137 CCAGAGTTGGGTGACCAAGATTTGCCCTGAGGATGAGGTTCTGGGCTCGAGGACACAGA 7196
QY 781 ccttactgactgagcc 798
DB 7197 CCTTACTGAGGTTGCC 7214
```

Search completed: January 8, 2002, 17:27:05  
Job time: 12716 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 8, 2002, 17:31:47 ; Search time 428.83 Seconds  
(without alignments)  
1999.219 Million cell updates/sec

Title: US-09-740-211-14\_COPY\_4000\_4999  
Perfect score: 1000  
Sequence: 1 gataaagtgatctgttgg.....cgcagctgcctgcagacat 1000

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1661242

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: N\_Geneseq\_1101:\*  
2: /SIDS2/gcgdata/geneseq/geneseqn/NA1980.DAT:\*  
3: /SIDS2/gcgdata/geneseq/geneseqn/NA1981.DAT:\*  
4: /SIDS2/gcgdata/geneseq/geneseqn/NA1982.DAT:\*  
5: /SIDS2/gcgdata/geneseq/geneseqn/NA1983.DAT:\*  
6: /SIDS2/gcgdata/geneseq/geneseqn/NA1984.DAT:\*  
7: /SIDS2/gcgdata/geneseq/geneseqn/NA1985.DAT:\*  
8: /SIDS2/gcgdata/geneseq/geneseqn/NA1986.DAT:\*  
9: /SIDS2/gcgdata/geneseq/geneseqn/NA1987.DAT:\*  
10: /SIDS2/gcgdata/geneseq/geneseqn/NA1988.DAT:\*  
11: /SIDS2/gcgdata/geneseq/geneseqn/NA1989.DAT:\*  
12: /SIDS2/gcgdata/geneseq/geneseqn/NA1990.DAT:\*  
13: /SIDS2/gcgdata/geneseq/geneseqn/NA1991.DAT:\*  
14: /SIDS2/gcgdata/geneseq/geneseqn/NA1992.DAT:\*  
15: /SIDS2/gcgdata/geneseq/geneseqn/NA1993.DAT:\*  
16: /SIDS2/gcgdata/geneseq/geneseqn/NA1994.DAT:\*  
17: /SIDS2/gcgdata/geneseq/geneseqn/NA1995.DAT:\*  
18: /SIDS2/gcgdata/geneseq/geneseqn/NA1996.DAT:\*  
19: /SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT:\*  
20: /SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT:\*  
21: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:\*  
22: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:\*  
23: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1000	100.0	4999	21	Recombinant adeno
2	1000	100.0	4999	22	Human factor VIII
3	954.4	95.4	11933	21	Recombinant adeno
4	954.4	95.4	11933	22	Human factor VIII
5	870.8	87.1	7944	22	Plasmodium falciparum
6	790	79.0	4275	10	DNA encoding 740 A
7	790	79.0	4629	20	Human factor VIII
8	790	79.0	4670	19	Human factor VIII
9	790	79.0	4882	19	Human factor VIII
10	790	79.0	4832	19	Human factor VIII
11	790	79.0	5035	18	Factor VIII-dB695-

12	790	79.0	5094	21	AAA49231	DNA construct HSO/
13	790	79.0	7056	15	AA066615	Sequence of human
14	790	79.0	8967	17	AA031031	Factor VIII full-1
15	790	79.0	8967	22	AA087526	Human factor VIII
16	790	79.0	8975	6	AA050054	Human factor VIII
17	790	79.0	8975	21	AA028604	Human full-length
18	790	79.0	9009	14	AA050185	Human factor VIII
19	790	79.0	9009	18	AA051548	Human factor VIII
20	790	79.0	9009	19	AA025810	Human factor VIII
21	790	79.0	9009	19	AA021212	Human factor VIII
22	790	79.0	9009	22	AA021162	Human factor VIII
23	790	79.0	9009	22	AA030508	Human factor VIII
24	790	79.0	9068	19	AA015359	Human factor VIII
25	790	79.0	9068	19	AA015359	Human factor VIII
26	790	79.0	9068	19	AA015359	Human factor VIII
27	790	79.0	9068	19	AA015359	Human factor VIII
28	790	79.0	9354	18	AA073164	Human factor III e
29	790	79.0	11846	20	AA082261	Beta-domain delete
30	790	79.0	12022	20	AA082260	CDNA encoding huma
31	790	79.0	12445	21	AA049232	Factor VIII protel
32	788.4	78.8	4616	9	AA081545	Factor VIII protei
33	788.4	78.8	4630	9	AA081544	Factor VIII protei
34	788.4	78.8	7059	9	AA081543	Factor VIII protei
35	788.4	78.8	8241	9	AA081539	Factor VIII protei
36	788.4	78.8	8241	9	AA081096	Factor VIII protei
37	788.4	78.8	9009	19	AA018884	Factor VIII protei
38	787	78.7	4272	9	AA080447	Factor VIII protei
39	787	78.7	4275	9	AA080446	Factor VIII protei
40	787	78.7	4545	9	AA080446	Factor VIII protei
41	787	78.7	7053	18	AA051357	Factor VIII:C codi
42	785.2	78.5	7440	7	AA060689	Sequence encoding
43	785.2	78.5	7440	16	AA076016	B-domain deleted F
44	764.4	76.4	7440	6	AA050375	DNA sequence encod
45	741	74.1	4373	20	AA082258	Beta-domain delete

## ALIGNMENTS

RESULT 1	
AD00122	standard; DNA; 4999 BP.
ID	AD00122
AC	AD00122
XX	31-JUL-2000 (first entry)
DE	Recombinant adeno associated vector construct, pV4.1cFdeltaB.
XX	Recombinant Adeno Associated Vector: rAAV; pV4.1cFdeltaB; promoter;
KW	human Factor VIII; hFVIII; EFTalpha; human elongation factor-1alpha;
KW	human growth hormone; hGH; ITR: inverted terminal repeat; haemophilia;
KW	gene therapy; ds.
XX	Adeno associated virus.
OS	Adeno associated virus.
XX	MO200023116-A1.
PN	MO200023116-A1.
XX	27-APR-2000.
PD	27-APR-2000.
XX	19-OCT-1999; 99WO-US24495.
PF	19-OCT-1999; 99WO-US24495.
XX	20-OCT-1998; 98US-0104994.
PR	24-MAR-1999; 99US-0125974.
PR	30-JUL-1999; 99US-0364862.
XX	(AVIC-) AVIGEN INC.
PA	(AVIC-) AVIGEN INC.
XX	Couto LB, Colosi PC;
PI	Couto LB, Colosi PC;
XX	WPI; 2000-339536/29.
DR	WPI; 2000-339536/29.
XX	New recombinant adenovirus-associated vector, useful for gene therapy
PT	New recombinant adenovirus-associated vector, useful for gene therapy

PT to treat hemophilia, comprises at least a portion of Factor VIII  
 PT operably linked to control sequence -  
 XX  
 XX Example 9; Fig 6; 92pp; English.

XX The present DNA sequence is a recombinant adenovirus-associated vector,  
 CC (RAV) construct pV4.1c8deltaB. This expression vector comprises the  
 CC promoter, first Intron (-573 to +985) of human elongation factor-1alpha  
 CC (EF1alpha) gene, human Factor VIII coding sequence (hFVIII) and a  
 CC polyadenylation signal from human growth hormone (hGH). This sequence is  
 CC inserted between the AAV inverted terminal repeat (ITR) regions. The  
 CC hFVIII coding region comprises the heavy chain gene segment with the  
 CC first 57 bp encoding the 19 amino acid signal peptide, A1 and A2 domains  
 CC and 5 amino acids from the C-terminal 85 amino acids of B domain and the A3,  
 CC C1 and C2 domains. Both the heavy and light chain segments are cloned  
 CC into the same plasmid separated by 42 nucleotides coding for 14 residues  
 CC of the B domain, that is deleted. This plasmid is operably linked to  
 CC control sequences, that directs the transcription and translation of the  
 CC Factor VIII gene. The adeno-associated viral vectors are used for gene  
 CC therapy to treat haemophilia. This method allows prolonged expression of  
 CC therapeutic levels of Factor VIII in vivo. The RAV are used for gene  
 CC therapy, because of their broad host range, safety profile and duration  
 CC of expression in the infected hosts.  
 CC  
 XX Sequence 4999 BP: 1375 A; 1154 C; 1155 G; 1318 T; 0 other;

Query Match 100.0%; Score 1000; DB 21; Length 4999;  
 Best Local Similarity 100.0%; Pred. No. 5, 6e-302;  
 Matches 1000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 gatcaagtgatgctgttgaccatgattatcaccgcatcaagaccagggtgccg 60  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 4000 gatcaagtgatgctgttgaccatgattatcaccgcatcaagaccagggtgccg 4059  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 61 tcagaagttctccagcctctacatctcagttatcatatgatagtttatggaa 120  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 4060 tcagaagttctccagcctctacatctcagttatcatatgatagtttatggaa 4119  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 121 gaatgacagatctcaggaatcctcactggaacctaaagtctcttggcaagt 180  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 4120 gaatgacagatctcaggaatcctcactggaacctaaagtctcttggcaagt 4179  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 181 ggattccttggaataaacaacatattttaaccctcaattatgctcgttaactcg 240  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 4180 ggattccttggaataaacaacatattttaaccctcaattatgctcgttaactcg 4239  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 241 ttggaccacatcatatagcatctcgacactcttcgcatgagttgagtgctgta 300  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 4240 ttggaccacatcatatagcatctcgacactcttcgcatgagttgagtgctgta 4299  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 301 tttaaatagttgacagatcgcatcttggaatggaataagaacatcagacagat 360  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 4300 tttaaatagttgacagatcgcatcttggaatggaataagaacatcagacagat 4359  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 361 tactgtctactacttaaccaataatgttgcacactgctcctcaaaagctcgact 420  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 4360 tactgtctactacttaaccaataatgttgcacactgctcctcaaaagctcgact 4419  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 421 tcacctccaagaggaggaatgactcgtgagactcagttgataatccaagaagtgcct 480  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 4420 tcacctccaagaggaggaatgactcgtgagactcagttgataatccaagaagtgcct 4479  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 481 gcaagtgcactccagaagaacatgaaagtcacagagtaactcagggagtaaatc 540  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 4480 gcaagtgcactccagaagaacatgaaagtcacagagtaactcagggagtaaatc 4539  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 541 tctgctaccacagatgataagagattcctcaatccagcagtcacaatgagtcacatca 600  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 4540 tctgctaccacagatgataagagattcctcaatccagcagtcacaatgagtcacatca 4599  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 601 gtgagctctcttttcggaagtggaagaaggtttttcagggaatacaagactcctt 660  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||

DB 4600 gtgagctctcttttcggaagtggaagaaggtttttcagggaatacaagactcctt 4659  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 661 cacacctgtgtgaactctctgaccaccacgcttactgactcactcctgaattacc 720  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 4660 cacacctgtgtgaactctctgaccaccacgcttactgactcactcctgaattacc 4719  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 721 ccagagttggtgacacgaattgacctgagatgaggtcttcgggctcggagacagga 780  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 4720 ccagagttggtgacacgaattgacctgagatgaggtcttcgggctcggagacagga 4779  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 781 cctctactgactcagaccataaaggaattatcttcatgcaatgagtggtgtt 840  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 4780 cctctactgactcagaccataaaggaattatcttcatgcaatgagtggtgtt 4839  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 841 ttgtgtggtgacacgaacccctagtgatgaggttgccaccctctctgagctc 900  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 4840 ttgtgtggtgacacgaacccctagtgatgaggttgccaccctctctgagctc 4899  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 901 gctcgtctactgagcgggagcgaacaaagtcgcccagcgcgggttgcggcg 960  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 4900 gctcgtctactgagcgggagcgaacaaagtcgcccagcgcgggttgcggcg 4959  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 961 cctcagtgagcagcagcgcagctcgtcgcagagcat 1000  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 4960 cctcagtgagcagcagcgcagctcgtcgcagagcat 4999  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||

# RESULT 2

AD08613  
 ID AAD08613 standard; DNA; 4999 BP.

AC AAD08613;  
 XX

DT 04-SEP-2001 (first entry)

XX Human factor VIII expressing RAV vector pV4.1c8-B partial sequence.

XX Recombinant adeno-associated viral vector; RAV; factor VIII; FVIII;  
 KW blood clotting disorder; gene therapy; haemophilia A; human;  
 KW pV4.1c8-B; ds.

XX Chimeric - Adeno associated virus.

OS Chimeric - Homo sapiens.

OS Chimeric - Unidentified.

PN W0200145510-A1.

PD 28-JUN-2001.

PF 21-DEC-2000; 2000WO-US34925.

PR 22-DEC-1999; 99US-0470618.

PA (AVIG-) AVIGEN INC.

PI Couto LB, Colosi PC, Qian X;

DR WPI; 2001-417955/44.

PT Treating blood clotting disorder, especially hemophilia in mammals, by  
 PT administering recombinant adeno-associated vectors which express blood  
 PT coagulation factor VIII -

XX Claim 19; Fig 6; 90pp; English.

XX The present invention relates to a method for treating a subject  
 CC suffering from a blood clotting disorder. The method comprises  
 CC administering a recombinant adeno-associated virion (RAV) comprising  
 CC a nucleotide sequence encoding the light chain of factor VIII and a  
 CC second recombinant adeno-associated virion comprising a nucleotide  
 CC sequence encoding the heavy chain of factor VIII. The RAV vector is  
 CC useful in gene therapy for treating haemophilia A in mammals, in



particular humans. The rAAV vector provides high level and long term expression of biologically active clotting factor VIII in vivo. The present sequence is pV4.1cF8-B vector without the plasmid backbone. It comprises human elongation factor 1 (EF1) promoter, B-domain deleted human factor VIII coding sequence, a poly A signal and AAV inverted terminal repeats (ITRs) at the ends. The vector encodes both the light and heavy chains of human factor VIII separated by 14 amino acids of the B domain.

Sequence 4999 BP; 1375 A; 1151 C; 1155 G; 1318 T; 0 other:

Query Match 100.0%; Score 1000; DB 22; Length 4999;  
Best Local Similarity 100.0%; Pred. No. 5,6e-302;  
Matches 1000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

0Y 1 gataaagtgatcgtgtgacccaatgttatttcaaggacccaagggtgccc 60  
DB 4000 gatcaagtgatcgtgtgacccaatgttatttcaaggacccaagggtgccc 4059  
0Y 61 tcaaaatttcacgctcacaatctcctcagttatcatatgatatgttgatgaa 120  
DB 4060 tcaaaatttcacgctcacaatctcctcagttatcatatgatatgttgatgaa 4119  
0Y 121 gaatggcagactatcaggaatattccatctgaacattatgtctcttctgcaatgt 180  
DB 4120 gaatggcagactatcaggaatattccatctgaacattatgtctcttctgcaatgt 4179  
0Y 181 ggatcattcgtggaataaacaacataatttaacctccaattatgtctgataatccg 240  
DB 4180 ggatcattcgtggaataaacaacataatttaacctccaattatgtctgataatccg 4239  
0Y 241 ttgacccaactaattatagatctgcagactcttcgataagtgatgagtcgtga 300  
DB 4240 ttgacccaactaattatagatctgcagactcttcgataagtgatgagtcgtga 4299  
0Y 301 ttttaaatgttcagatcagatcttggaatgagagataaagataatcagatcagat 360  
DB 4300 ttttaaatgttcagatcagatcttggaatgagagataaagataatcagatcagat 4359  
0Y 361 tactgttcatcctacttaccataatgtttgacacactggtctcttcaaaagctcagat 420  
DB 4360 tactgttcatcctacttaccataatgtttgacacactggtctcttcaaaagctcagat 4419  
0Y 421 tccaccccaaggagagagaaatgacctgagagactcagtgataatccaaaagtggtc 480  
DB 4420 tccaccccaaggagagagaaatgacctgagagactcagtgataatccaaaagtggtc 4479  
0Y 481 gcaagtgacttcagaagaacatgaatgaacagagtaactactcagaggaatgaatc 540  
DB 4480 gcaagtgacttcagaagaacatgaatgaacagagtaactactcagaggaatgaatc 4539  
0Y 541 tctgttaccagcagatgatgtgaagaggttccatctcagcagatcagaatgagcctaca 600  
DB 4540 tctgttaccagcagatgatgtgaagaggttccatctcagcagatcagaatgagcctaca 4599  
0Y 601 gtgagactctttttcagaatgagcaagatggttttccaggggaatcaagactcctt 660  
DB 4600 gtgagactctttttcagaatgagcaagatggttttccaggggaatcaagactcctt 4659  
0Y 661 caccacttggtgaactctctagaccacagcttactcgtactcctcgaattcacc 720  
DB 4660 caccacttggtgaactctctagaccacagcttactcgtactcctcgaattcacc 4719  
0Y 721 ccagagttgtgacacagagagcccgagagatgaggttctggtcgtgagcacagga 780  
DB 4720 ccagagttgtgacacagagagcccgagagatgaggttctggtcgtgagcacagga 4779  
0Y 781 cctctactgacgagcccaataaagaaatttattcattgacatagtggttggtt 840  
DB 4780 cctctactgacgagcccaataaagaaatttattcattgacatagtggttggtt 4839  
0Y 841 ttgtgtgagggcgacaggaacccctagtgatgttgacactccctctctgagcgtc 900

DB 4840 ttgtgtgagggcgacaggaacccctagtgatgttgacactccctctctgagcgtc 4899  
0Y 901 gctcgtcactgagcgcgagcagacaaagtcgacgacgcgagcttggccgggagc 960  
DB 4900 gctcgtcactgagcgcgagcagacaaagtcgacgacgcgagcttggccgggagc 4959  
0Y 961 cctcagtgagagagcagcgcgacagctgctgcagagat 1000  
DB 4960 cctcagtgagagagcagcgcgacagctgctgcagagat 4999

RESULT 3  
AAD00121 standard; DNA: 11933 BP.  
ID AAD00121  
AC AAD00121;  
XX 31-JUL-2000 (first entry)  
DE Recombinant adeno associated vector construct, pAAV-F8-1.  
XX Recombinant Adeno Associated Vector: rAAV; pAAV-F8-1; human Factor VIII;  
KW hFVIII; hNF-3 albumin promoter; human elongation factor-1alpha; EF1alpha;  
KW human growth hormone; hGH; Inverted terminal repeat; ITR; haemophilia;  
KW gene therapy: ds.  
XX Adeno associated virus.  
XX WO200023116-A1.  
PN 27-APR-2000.  
PD 19-OCT-1999; 99WO-US24495.  
XX 20-OCT-1998; 98US-0104994.  
PR 24-MAR-1999; 98US-0125974.  
PR 30-JUL-1999; 99US-0364862.  
XX (AVIG-) AVIGEN INC.  
XX Couto LB, Colosi PC;  
PI WPI: 2000-339536/29.  
DR New recombinant adenovirus-associated vector, useful for gene therapy  
PT to treat hemophilia, comprises at least a portion of Factor VIII  
PI operably linked to control sequence  
XX Example 2; Fig 5; 92pp; English.

PS The present DNA sequence is a recombinant adenovirus-associated vector, (rAAV) construct, pAAV-F8-1. This expression vector comprises the hNF-3 (rAAV) promoter, the first intron (-573 to +985) of human elongation factor-1alpha (EF1alpha) gene, human Factor VIII coding sequence (hFVIII) and polyadenylation signal from human growth hormone (hGH). This sequence is inserted between the AAV inverted terminal repeat (ITR) regions. The hFVIII coding region comprises the heavy chain segment with the first 57 bp encoding the 19 amino acid signal peptide, A1 and A2 domains and 5 amino acids from the C-terminal 85 amino acids of B domain and the A3, segment comprises the C-terminal 14 amino acids of B domain and the A3, C1 and C2 domains. Both the heavy and light chain segments are cloned into the same plasmid separated by 42 nucleotides coding for 14 residues of the B domain, that is deleted. This plasmid is operably linked to the control sequences, that directs the transcription and translation of the Factor VIII gene. The adeno-associated viral vectors are used for gene therapy to treat haemophilia. This method allows prolonged expression of therapeutic levels of Factor VIII in vivo. The rAAV are used for gene therapy, because of their broad host range, safety profile and duration of expression in the infected hosts.

Sequence 11933 BP; 3258 A; 2818 C; 2717 G; 3140 T; 0 other;

Query Match 95.4%; Score 954.4; DB 21; Length 11933;  
 Best Local Similarity 97.9%; Pred. No. 1.7e-287;  
 Matches 979; Conservative 0; Mismatches 16; Indels 5; Gaps 1;

1 gataaagttgactgtgttgacacaaatgattatcaagcgaacaaaggggtgccg 60  
 |||||||  
 Db 3975 gatcaagggtgactgtgttgacacaaatgattatcaagcgaacaaaggggtgccg 4034  
 QY 61 tcgaagttctccagcctcatctccagttatcatatgatatgtcttgatggaa 120  
 |||||||  
 Db 4035 tcgaagttctccagcctcatctccagttatcatatgatatgtcttgatggaa 4094  
 QY 121 gaagtgacagacttctcgaagaatctcaactggaacttaagtctcttggcaatg 180  
 |||||||  
 Db 4095 gaagtgacagacttctcgaagaatctcaactggaacttaagtctcttggcaatg 4154  
 QY 181 ggattcatctgggataaacaacataattttaaccctccaatatgtctcgatataccg 240  
 |||||||  
 Db 4155 ggattcatctgggataaacaacataattttaaccctccaatatgtctcgatataccg 4214  
 QY 241 ttggacccaactattatagatctcgacgaactcttcgatggatgatatggtctgta 300  
 |||||||  
 Db 4215 ttggacccaactattatagatctcgacgaactcttcgatggatgatatggtctgta 4274  
 QY 301 tttaaatagtgcagcatgcatattgggaatggagaaagaacataatagatgcagat 360  
 |||||||  
 Db 4275 tttaaatagtgcagcatgcatattgggaatggagaaagaacataatagatgcagat 4334  
 QY 361 tactgtctatctacttaccataatgttgcacactgtgtctcttaaaagctcgact 420  
 |||||||  
 Db 4335 tactgtctatctacttaccataatgttgcacactgtgtctcttaaaagctcgact 4394  
 QY 421 tcaactcagaaggaggaatgactcctgagacactcaggttgataatcaaaaggtgct 480  
 |||||||  
 Db 4395 tcaactcagaaggaggaatgactcctgagacactcaggttgataatcaaaaggtgct 4454  
 QY 481 gcaagtgacttccagaagaacataatgaatgaacagagtaactcagaaggagtaaatc 540  
 |||||||  
 Db 4455 gcaagtgacttccagaagaacataatgaatgaacagagtaactcagaaggagtaaatc 4514  
 QY 541 tctgcttaccagcatgtatgtgaagaatctctcatctccagcggatgaatgacatac 600  
 |||||||  
 Db 4515 tctgcttaccagcatgtatgtgaagaatctctcatctccagcggatgaatgacatac 4574  
 QY 601 gtggaactcctttttcagaatgtaagaatgttttcagaagaaatcaaacctct 660  
 |||||||  
 Db 4575 gtggaactcctttttcagaatgtaagaatgttttcagaagaaatcaaacctct 4634  
 QY 661 caccctgtgtgaactctcagacccaacgttactgaactcgtactctcgaattacc 720  
 |||||||  
 Db 4635 caccctgtgtgaactctcagacccaacgttactgaactcgtactctcgaattacc 4694  
 QY 721 ccaaggttggtgacacagatgtccctgagatgtaggtcttgggtgcgaagcacaaga 780  
 |||||||  
 Db 4695 ccaaggttggtgacacagatgtccctgagatgtaggtcttgggtgcgaagcacaaga 4754  
 QY 781 cctactgactcagacctataaagaatattattatctgcaaatgtgtgtgatt 840  
 |||||||  
 Db 4755 cctactgactcagacctataaagaatattattatctgcaaatgtgtgtgatt 4809  
 QY 841 ttgtgtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 900  
 |||||||  
 Db 4810 ttgtgtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 4869  
 QY 901 gctcgtcactgagc 960  
 |||||||  
 Db 4870 gctcgtcactgagc 4929  
 QY 961 cctcagttgagc 1000  
 |||||||  
 Db 4930 cctcagttgagc 4969

# RESULT 4

ID AAD08612 standard; DNA; 11933 BP.

AC AAD08612;

DT 04-SEP-2001 (first entry)

DE Human factor VIII expressing RAAV vector PAAV-F8-1 partial sequence.

RE Recombinant adeno-associated viral vector; RAAV; factor VIII; FVIII;

KW blood clotting disorder; gene therapy; haemophilia A; human;

KW PAAV-F8-1; ds.

OS Chimeric - Adeno associated virus.

OS Chimeric - Mus sp.

OS Chimeric - Homo sapiens.

OS Chimeric - Oryctolagus cuniculus.

PN W020014510-A1.

XX 28-JUN-2001.

PF 21-DEC-2000; 2000WO-US34925.

PR 22-DEC-1999; 99US-0470618.

XX (AVIG-) AVIGEN INC.

XX Couto LB, Colosi PC, Qian X;

PI WPI; 2001-417955/44.

DR Treating blood clotting disorder, especially hemophilia in mammals, by

XX administering recombinant adeno-associated vectors which express blood

PT coagulation factor VIII

XX Claim 18; Fig 5; 90pp; English.

PS The present invention relates to a method for treating a subject

XX suffering from a blood clotting disorder. The method comprises

CC administering a recombinant adeno-associated virion (RAAV) comprising

CC a nucleotide sequence encoding the light chain of factor VIII and a

CC second recombinant adeno-associated virion comprising a nucleotide

CC sequence encoding the heavy chain of factor VIII. The RAAV vector is

CC useful in gene therapy for treating haemophilia A in mammals, in

CC particular humans. The RAAV vector provides high level and long term

CC expression of biologically active clotting factor VIII in vivo.

CC The present sequence is PAAV-F8-1 vector without the plasmid

CC backbone. It comprises HNF-3, mouse albumin promoter, a synthetic intron

CC (I96) intron sequences, B-domain deleted human factor VIII coding

CC sequence, poly A signal based on rabbit beta-globin sequence and AAV

CC inverted terminal repeats (ITRs) at the ends. The vector encodes both the

CC light and heavy chains of human factor VIII.

XX Sequence 11933 BP; 3258 A; 2818 C; 2717 G; 3140 T; 0 other;

## Query Match

Best Local Similarity 95.4%; Score 954.4; DB 22; Length 11933;  
 Matches 979; Conservative 0; Mismatches 16; Indels 5; Gaps 1;

1 gataaagttgactgtgttgacacaaatgattatcaagcgaacaaaggggtgccg 60  
 |||||||  
 Db 3975 gatcaagggtgactgtgttgacacaaatgattatcaagcgaacaaaggggtgccg 4034  
 QY 61 tcgaagttctccagcctcatctccagttatcatatgatatgtcttgatggaa 120  
 |||||||  
 Db 4035 tcgaagttctccagcctcatctccagttatcatatgatatgtcttgatggaa 4094  
 QY 121 gaagtgacagacttctcgaagaatctcaactggaacttaagtctcttggcaatg 180

```

Db      4095 gaagtgagcgaactatcgaggaattccactggaattcgaatgcttcttgcgaatgt 4154
          |||
Oy      181 ggaattcatctgggataaaacacaaatatttttaacctccaatatttgcgatacatcgc 240
          |||
Db      4155 ggaattcatctgggataaaacacaaatatttttaacctccaatatttgcgatacatcgc 4214
          |||
Oy      241 ttgcgacccaactcatatagatctgcgacacactctgcgatagagtgatgagctgtga 300
          |||
Db      4215 ttgcgacccaactcatatagatctgcgacacactctgcgatagagtgatgagctgtga 4274
          |||
Oy      301 tttaaatagttgacgactgacccattgggaatggagtaagaataatacgaatgacagat 360
          |||
Db      4275 tttaaatagttgacgactgacccattgggaatggagtaagaataatacgaatgacagat 4334
          |||
Oy      361 tactgcttacccttacttacttaccataatgtcttgcacacttgctccttcaaaagctcgc 420
          |||
Db      4335 tactgcttacccttacttacttaccataatgtcttgcacacttgctccttcaaaagctcgc 4394
          |||
Oy      421 tcacctccaaggagaggaataatgctgaggaacttcgaagtggaataatccaaagatgagct 480
          |||
Db      4395 tcacctccaaggagaggaataatgctgaggaacttcgaagtggaataatccaaagatgagct 4454
          |||
Oy      481 gcaagtggaacttcagaagaacaaatgaagaatcagaagagtaactactcagaaggagtaaatc 540
          |||
Db      4455 gcaagtggaacttcagaagaacaaatgaagaatcagaagagtaactactcagaaggagtaaatc 4514
          |||
Oy      541 tctgttaccagcatgtatgtgaagaggttcctcatctccagagtgcaaatgagccatca 600
          |||
Db      4515 tctgttaccagcatgtatgtgaagaggttcctcatctccagagtgcaaatgagccatca 4574
          |||
Oy      601 gtagagctcttttttgaagatgagcaagaagtatttttcagaagaaatcaaatctcctt 660
          |||
Db      4575 gtagagctctctttttgaagatgagcaagaagtatttttcagaagaaatcaaatctcctt 4634
          |||
Oy      661 caacactgtgtgaaactctctagaccacgcttactgacttgcactcttgcgaataacc 720
          |||
Db      4635 caacactgtgtgaaactctctagaccacgcttactgacttgcactcttgcgaataacc 4694
          |||
Oy      721 ccagagttggtgacccagatgctcctagagatgagagttctgagcttgagagacaga 780
          |||
Db      4695 ccagagttggtgacccagatgctcctagagatgagagttctgagcttgagagacaga 4754
          |||
Oy      781 cctctactgactgagcactaataaagaaatttatttcattgcaatagtgtgtgtt 840
          |||
Db      4755 cctctactgactgagcactaataaagaaatttatttcattgcaatagtgtgtgtt 4809
          |||
Oy      841 ttgtgtgtgagcgagcaggaacccctagtgatgagttgagcactcctctctgcgcgtc 900
          |||
Db      4810 ttgtgtgtgagcgagcaggaacccctagtgatgagttgagcactcctctctgcgcgtc 4869
          |||
Oy      901 gctgcgactactgagcgcgagcgaacaaaggtcgccgaagcccgagcttgcgcgcgcgc 960
          |||
Db      4870 gctgcgactactgagcgcgagcgaacaaaggtcgccgaagcccgagcttgcgcgcgcgc 4929
          |||
Oy      961 cctcagtgagcgagcgagcgcgagcgtgcgctgcgagacat 1000
          |||
Db      4930 cctcagtgagcgagcgagcgcgagcgtgcgctgcgagacat 4969
          |||

```

## RESULT 5

AAAF84647 standard; DNA: 7944 BP.

ID AAF84647

AC AAF84647

XX 29-JUN-2001 (first entry)

XX Plasmid DL26 encoding human B-domain deleted factor VIII.

XX Adeno-associated virus vector; B-domain; factor VIII; haemophilia A;

XX coagulation disorder; ss.

```

OS Synthetic.
OS Homo sapiens.
OS Hepatitis B virus.
XX
FH key Location/Qualifiers
FT 1..146
FT misc_feature
FT /tag= a
FT /note= "Inverted terminal repeat"
FT enhancer
FT 150..278
FT /tag= b
FT /note= "hepatitis B virus Enh enhancer"
FT CDS
FT 420..4835
FT /tag= c
FT /note= "human B-domain deleted factor VIII"
FT polyA_signal
FT 4840..4914
FT /tag= d
FT /note= "TK polyA sequence"
FT misc_feature
FT 4916..5084
FT /tag= e
FT /note= "Inverted terminal repeat"
XX
XX W0200127303-A1.
XX 19-APR-2001.
XX 12-OCT-2000; 2000MO-US28221.
XX 12-OCT-1999; 99US-0158780.
XX (DUNC-) UNIV NORTH CAROLINA.
XX Walsh CE, Chao H, Burstein H, Lynch CM, Stepan AM, Munson K;
XX WPI: 2001-273781/28.
XX P-PSDB; AAB67959.
XX
XX New recombinant adeno-associated virus vector, useful for treating
XX haemophilia A, comprises heterologous nucleotide sequence encoding
XX B-domain deleted human factor VIII operably linked with liver-preferred
XX expression control element -
XX
XX Claim 64; Fig 1; 87pp; English.
XX
XX The specification describes a recombinant adeno-associated virus (rAAV)
XX vector. The vector comprises a heterologous nucleotide sequence
XX encoding B-domain deleted factor VIII operably linked with at least one
XX enhancer and at least one promoter. The method results in the production
XX of high titer rAAV vector stocks carrying the B-domain deleted factor
XX CC VIII transgenes and expression cassettes, which generate adequate titers
XX of virus for in vivo administration. The recombinant vectors are useful
XX CC for treating haemophilia A, where the liver expresses the encoded
XX CC B-domain deleted factor VIII, which is secreted into the blood. They are
XX CC also useful for the treatment of other coagulation disorders. The
XX CC present sequence encodes a B-domain deleted factor VIII.
XX
XX Sequence 7944 BP; 2142 A; 1902 C; 1909 G; 1991 T; 0 other;
SQ

```

## Query Match

87.1%; Score 870.8; DB 22; Length 7944;

Best Local Similarity 92.7%; Pred. No. 1.8e-261;

Matches 933; Conservative 0; Mismatches 52; Indels 21; Gaps 1;

```

Oy      1 gatcaagtgatctgttggacaaatgattatccaagcagcaaacacaggtgcccgc 60
          |||
Db      4046 gatcaagtgatctgttggacaaatgattatccaagcagcaaacacaggtgcccgc 4105
          |||
Oy      61 tcagaagttccacgctctacatctcgaattatcatatglatagcttgatggaa 120
          |||
Db      4106 tcagaagttccacgctctacatctcgaattatcatatglatagcttgatggaa 4165
          |||
Oy      121 gaagtgcagactatcgaggaatccactggaaccttaatgtctcttggcaatgt 180
          |||
Db      4166 gaagtgcagactatcgaggaatccactggaaccttaatgtctcttggcaatgt 4225
          |||

```

OY	181	ggatccatctgaggtaaacacaataatttaacccctcaattatgtctcgatacatccg	240
Db	4226	ggatccatctgaggtaaaaaacacabatttttaacctccatctaatatgtctcgatacatccg	4287
OY	241	tttgaccacacatcatatagacatctcgagacactcttcgcattgagttgattggtcgtga	300
Db	4286	tttgaccacacatcatatagacatctcgagacactcttcgcattgagttgattggtcgtga	4345
OY	301	tttaataagtttcgaagatcccatctggagaatgagagtaagaacaaatcatcatgacaaat	360
Db	4346	tttaataagtttcgaagatcccatctggagaatgagagtaagaacaaatcatcatgacaaat	4405
OY	361	taacgtctcatctactattacaataatgtttggccacctgtctctctccaagaatccgat	420
Db	4406	taacgtctcatctactattacaataatgtttggccacctgtctctctccaagaatccgat	4465
OY	421	tcaactccaaaggagagtaatgtcctctgagacccctcaagttaataatccaaagaattgct	480
Db	4466	tcaactccaaaggagagtaatgtcctctgagacccctcaagttaataatccaaagaattgct	4525
OY	481	gcaagtggaactccagaagaacaatgaaagctcaacagatctaactactcaaggagtaaatc	540
Db	4526	gcaagtggaactccagaagaacaatgaaagctcaacagatctaactactcaaggagtaaatc	4585
OY	541	tctgtcttaaccaatgatactgtaagaagggttcctccatctccagcagtcagaattggtccatca	600
Db	4586	tctgtcttaaccaatgatactgtaagaagggttcctccatctccagcagtcagaattggtccatca	4645
OY	601	gttggaactccttttttcaagaatgagcaagtaaaagtctttcaaggagaaatcaagaactcctt	660
Db	4646	gttggaactccttttttcaagaatgagcaagtaaaagtctttcaaggagaaatcaagaactcctt	4705
OY	661	caacacctgtgtgaaactcctcagaaccacacccgttactcgaatcgtcactcctcgaaatcaacc	720
Db	4706	caacacctgtgtgaaactcctcagaaccacacccgttactcgaatcgtcactcctcgaaatcaacc	4765
OY	721	ccaaggttgggttcgacagaatctgcctctgagatgagatgagttctctgagctgagagcaacagga	780
Db	4766	ccaaggttgggttcgacagaatctgcctctgagatgagatgagttctctgagctgagagcaacagga	4825
OY	781	cctcaactgagatcgagc-----ctataaagaaattatatttc	819
Db	4826	cctcaactgagatcgagatctctctctgagaggaatcggccaataaagaacagataaacaac	4885
OY	820	attgcgaatagtgatgtgattttttgtgtgacccgacgaagaaacccctagtgatgtg	879
Db	4886	gcaacgggtgtgtgtgcgtgtgttttcgataccagatcctagaaacccctagtgatgtg	4945
OY	880	ccaactccctctcggcgcgttcgtctcaactgagcccgagcgacaacaaagtctgcgcagac	939
Db	4946	ccaactccctctcggcgcgtctcgtctcaactgagcccgagcgacaacaaagtctgcgcagac	5005
OY	940	gcccgcgttttgcggcgcgccctcaagtgaaacggagcgagcgcgag	985
Db	5006	gggcgcaccttgttgcgcgcgctcaagtgaagcgagcgagcgcgag	5051

RESULT	6
AA90654	
ID	AA90654 standard; DNA: 4275 BP.
XX	
AC	AA90654;
XX	
DT	26-JUN-1990 (first entry)
XX	
DE	DNA encoding 740 Arg-1649 Glu human Factor VIII:C.
XX	
KW	Human Factor VIII:C; Ad.RE.neo; 740 Arg-1649 Glu Factor VIII:C;
KW	haemophilias A.
XX	
OS	Homo sapiens.
XX	

FH	Key	Location/Qualifiers
FT	CDS	1..4275
FT		/+tag= a
XX		
PN		EP306966-A.
XX		
PD		15-MAR-1989.
XX		
PF		09-SEP-1988; 88EP-0114769.
XX		
PR		08-APR-1988; 88JP-0085454.
XX		
PA		(KAGA ) CHEMO-SERO-THERAP.
PA		(TEIJU ) TEIJUN LTD.
XX		
PI		Sugiyama T., Masuda K., Tajima Y., Yonemura H;
DR		WPI: 1989-078467/11.
XX		P-PSDB; AAP91165.
PT		Prod'n. of recombinant human Factor-VIII-C -
PT		using animal cells transformed with a vector cong.
XX		Factor VIII:C and a promoter
XX		
PS		Fig 1(1) - 1(13); ; 32pp; English.
CC		
CC		When translated, Arg-740 of the carboxyl terminus of the H chain is
CC		directly bonded by a peptide bond to Glu-1649 of the amino terminus of
CC		L chain. It is used to transform animal cells so that they produce
CC		human Factor VIII:C. A pref'd. expression vector is plasmid Ad.RE.neo.
CC		The transformation can constantly at least one promoter upstream of AN90654.
CC		VIII:C in high yield on a commercial scale. The human Factor VIII:C so
CC		produced is considered to corresp. to the smallest species of active and
CC		inactive Factor VIII:C molecules in the human blood plasma. It is useful
XX		for treating haemophilia A patients.
XX		
SQ		Sequence 4275 BP; 1245 A; 941 C; 945 G; 1144 T; 0 other;
<hr/>		
	Query Match	79.0%; Score 790; DB 10; Length 4275;
	Best Local Similarity	100.0%; Pred. No. 2.4e-236;
	Matches 790; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
OY	1	gaccaagggtgagctgttgtagcaccgaatgatatttcagcgcatcaagaaccaggtgcggc 60
Db	3486	gaccaagtgagactctgttgcaccatgatacttcaagcattcaagaccagaggtgcggc 60
OY	61	tccgaagtctccagacctbaactctctaagtattatcatcagtgtaagtctgtatggaa 120
Db	3546	tccaagaattctccagacctbaactctctaagtattatcatcagtgtaagtctgtatggaa 120
OY	121	gaagtgcagacttatcggaggaattcccaacttgaaccttaatggtcctctttgacaatgt 180
Db	3606	gaagtgcagacttatcggaggaattcccaacttgaaccttaatggtcctctttgacaatgt 180
OY	181	ggatcattctgggataaacaacaatatattttaacctccaattatgttgtatatacccg 240
Db	3666	ggatcattctgggataaacaacaatatattttaacctccaattatgttgtatatacccg 240
OY	241	tttgaccaccaattatctatgacttcgagacacctcttcgcatggtgagtgtatggctgtga 300
Db	3726	tttgaccaccaattatctatgacttcgagacacctcttcgcatggtgagtgtatggctgtga 300
OY	301	tttaaatagtgtcagcaagcatctgggaatggagagtaaagcaatatcagatgcacagat 360
Db	3786	tttaaatagtgtcagcaagcatctgggaatggagagtaaagcaatatcagatgcacagat 360
OY	361	tacggttcacaccttaattaccaaatagtttgccaacctgtctccttcaaaaagctcagat 420
Db	3846	tacggttcacaccttaattaccaaatagtttgccaacctgtctccttcaaaaagctcagat 420
OY	421	tcaactccaaggaggagtaatgcttgtagaaccttcagggtgaataatccaaagaagtgcct 480



OS	Homo sapiens.
SS	Synthetic.
PN	MO9812207-AI.
PX	
PD	26-MAR-1998.
XX	
PF	18-SEP-1997; 97WO-US16639.
PR	20-SEP-1996; 96US-0717294.
XX	(GEHO ) GEN HOSPITAL CORP.
PA	
PI	Haas J, Seed B;
XX	
DR	WPI; 1998-217200/19.
XX	
PT	New synthetic eukaryotic gene(s) - in which non-preferred or less preferred codon(s) are replaced to provide high level expression in mammalian cell(s)
PP	
XX	
PS	
XX	Example 3; Fig 12; 92PP; English.
CC	This gene codes for a human Factor-VIII protein that lacks the central B domain (amino acids 760-1639) of the native protein. In a novel, claimed synthetic gene (see AAV23288), non-preferred or less preferred codons of the native gene are replaced by codons favored by highly expressed human genes to provide high-level expression in mammalian cells. The synthetic gene was assembled from 29 pairs of oligonucleotides (see AAV23340-97) which served as PCR templates. Synthetic genes of the invention (see also AAV23389-91) are used for production of recombinant proteins in mammalian cells at levels of at least 500% of those obtained using the natural genes. They can also be used in gene therapy.
XX	
SO	Sequence 4670 BP; 1325 A; 1072 C; 1036 G; 1227 T; 0 U.

	Query Match	Score 790:	DB 19:	Length 4670:
Best Local Similarity	79.08;			
Matches 793; Conservative	99.48;	Pred. No. 2.5e-236;		

QY		1	gatcaagtgagatctgtttggacccaatgatattcaagcgtcaacgaaccagggtgccg	60
Db	3653		gatcaagtgagatctgtttggacccaatgatattcaagcgtcaacgaaccagggtgccg <td>3712</td>	3712
QY	61		tcaagaattcccaagcctctacacvtctaaatttcaatcatgatagctctataggaa	120
Db	3713		tcaagaattcccaagcctctacacvtctaaatttcaatcatgatagctctataggaa	3772
QY	121		gaagtggcagactatcgaagaaattccactcgtgaaccttaagtctctcttggcaagt	180
Db	3773		gaagtggcagactatcgaagaaattccactcgtgaaccttaagtctctcttggcaagt	3832
QY	181		ggatctactcggagataaaacaacatattttaaccctccaattatctgtcgaataatcg	240
Db	3833		ggatctactcggagataaaacaacatattttaaccctccaattatctgtcgaataatcg	3892
QY	241		tttcgccaacacatataatgaatcttcgacgaactcttcgacgatgaaattgatagtgctgga	300
Db	3893		tttcgccaacacatataatgaatcttcgacgaactcttcgacgatgaaattgatagtgctgga	3952
QY	301		tttaaatagttgcagatgcacatcgtggagatgagagaaagcaatcgaatgcacagat	360
Db	3953		tttaaatagttgcagatgcacatcgtggagatgagagaaagcaatcgaatgcacagat	4012
QY	361		tactgctcatctcaactttaccaatatgtttgcaccctggctctcttaaaagctcagat	420
Db	4013		tactgctcatctcaactttaccaatatgtttgcaccctggctctcttaaaagctcagat	4072
QY	421		tcaacctcaaggagagagtaattgctcggagaccttcaggaaatccaaaagggtgct	480
Db	4073		tcaacctcaaggagagagtaattgctcggagaccttcaggaaatccaaaagggtgct	4132

QY	481	gcacagtgcgactccagaagacatgaaagtcacagagctactactccagggttaaatc	540
Db	4133	gcacagtgcgactccagaagacatgaaagtcacagagctactactccagggttaaatc	4192
QY	541	tctgcgtaccagacatgtagtgcgaaggagttccatctctccagagtcgaatgtgccatca	600
Db	4193	tctgcgtaccagacatgtagtgcgaaggagttccatctctccagagtcgaatgtgccatca	4252
QY	601	gtgcgactcctctttcttcgaagtccaagtcaaaagctttctcaggagaaatccaagactcct	660
Db	4253	gtgcgactcctctttcttcgaagtccaagtcaaaagctttctcaggagaaatccaagactcct	4312
QY	661	cacacctgtgttgtaactctcttagaccacacggtttactgcgtactctcttgaaattacc	720
Db	4313	cacacctgtgttgtaactctcttagaccacacggtttactgcgtactctcttgaaattacc	4372
QY	721	ccagagttggtgtgcacacagatctcctcttgaggttggaagttctgggctgcgagacagga	780
Db	4373	ccagagttggtgtgcacacagatctcctcttgaggttggaagttctgggctgcgagacagga	4432
QY	781	cctctactgcactgcagcc	798
Db	4433	cctctactgcaggtgagc	4450

RESULT	9
AAV19581	
XX	AAV19581 standard; cDNA; 4832 BP
ID	
XX	
AC	AAV19581;
XX	
DT	06-AUG-1998 (first entry)
XX	

	deleted	SQN deletion	CDNA sequence
1	.....	.....	.....
2	.....	.....	.....
3	.....	.....	.....
4	.....	.....	.....
5	.....	.....	.....
6	.....	.....	.....
7	.....	.....	.....
8	.....	.....	.....
9	.....	.....	.....
10	.....	.....	.....
11	.....	.....	.....
12	.....	.....	.....
13	.....	.....	.....
14	.....	.....	.....
15	.....	.....	.....
16	.....	.....	.....
17	.....	.....	.....
18	.....	.....	.....
19	.....	.....	.....
20	.....	.....	.....
21	.....	.....	.....
22	.....	.....	.....
23	.....	.....	.....
24	.....	.....	.....
25	.....	.....	.....
26	.....	.....	.....
27	.....	.....	.....
28	.....	.....	.....
29	.....	.....	.....
30	.....	.....	.....
31	.....	.....	.....
32	.....	.....	.....
33	.....	.....	.....
34	.....	.....	.....
35	.....	.....	.....
36	.....	.....	.....
37	.....	.....	.....
38	.....	.....	.....
39	.....	.....	.....
40	.....	.....	.....
41	.....	.....	.....
42	.....	.....	.....
43	.....	.....	.....
44	.....	.....	.....
45	.....	.....	.....
46	.....	.....	.....
47	.....	.....	.....
48	.....	.....	.....
49	.....	.....	.....
50	.....	.....	.....
51	.....	.....	.....
52	.....	.....	.....
53	.....	.....	.....
54	.....	.....	.....
55	.....	.....	.....
56	.....	.....	.....
57	.....	.....	.....
58	.....	.....	.....
59	.....	.....	.....
60	.....	.....	.....
61	.....	.....	.....
62	.....	.....	.....
63	.....	.....	.....
64	.....	.....	.....
65	.....	.....	.....
66	.....	.....	.....
67	.....	.....	.....
68	.....	.....	.....
69	.....	.....	.....
70	.....	.....	.....
71	.....	.....	.....
72	.....	.....	.....
73	.....	.....	.....
74	.....	.....	.....
75	.....	.....	.....
76	.....	.....	.....
77	.....	.....	.....
78	.....	.....	.....
79	.....	.....	.....
80	.....	.....	.....
81	.....	.....	.....
82	.....	.....	.....
83	.....	.....	.....
84	.....	.....	.....
85	.....	.....	.....
86	.....	.....	.....
87	.....	.....	.....
88	.....	.....	.....
89	.....	.....	.....
90	.....	.....	.....
91	.....	.....	.....
92	.....	.....	.....
93	.....	.....	.....
94	.....	.....	.....
95	.....	.....	.....
96	.....	.....	.....
97	.....	.....	.....
98	.....	.....	.....
99	.....	.....	.....
100	.....	.....	.....

KW replication defective; recombinant retrovirus; RRV; therapeutic protein;  
 KW haemophilia; thrombosis; hypercoagulable disorder; liver disease; human;  
 KW hepatitis; thalassaemia; phenylketonuria; Leach-Nathan syndrome; diabetes;  
 KW cystic fibrosis; Duchenne's Muscular Dystrophy; hypercholesterolemia;  
 KW hypoparathyroidism; adenosine deaminase deficiency; HIV infection; anaemia;  
 KW Gaucher's syndrome; high blood pressure; Alzheimer's disease; autoimmune  
 KW inflammatory disease; factor VIIIr55.

OS	Homo sapiens.	
XX		
PPN	W09800541-A2.	
XXX		
PPD	08-JAN-1998.	
PEP		
XXF		
XX	02-JUL-1997;	97WO-US11784
XX		
XX		
04-JUN-1997;		97US-0869309
03-JUL-1996;		96US-0645601
ROR		
13-AUG-1996;		96US-0696581.
XX		

(CHR ) CHIRON CORP

WPI: 1998-086966/08.  
P-PSDB: AAM46246.

administered to provide long term systemic expression of therapeutic protein in blood, useful in, e.g. treating hyper-coagulable disorders

Example 28; Pages 210-213; 272pp; English.

deleted SQN deletion protein of human

CC factor VIII. This is used in the construction of recombinant retroviral  
 CC vectors expressing human factor VIII. The invention provides the  
 CC preparation of replication defective recombinant retrovirus (RV)  
 CC expressing a therapeutic protein. The RV preparation is resistant to  
 CC degradation by human complement and is capable of inducing long term  
 CC systemic expression of the therapeutic protein when administered  
 CC intravenously to a human. The long term systemic expression results in a  
 CC measurable level of the therapeutic protein being produced in the blood  
 CC of the human for a period of at least 30 days after the administration of  
 CC the RV vector preparation. RV's can be used for in vivo delivery of  
 CC therapeutic protein to treat, e.g., haemophilia A, haemophilia B,  
 CC thrombosis, hypercoagulable disorders, liver diseases such as hepatitis,  
 CC disorders such as thalassemia, phenylketonuria, Lesch-Nyhan syndrome,  
 CC severe combined immunodeficiency (SCID), cystic fibrosis, Duchenne's  
 CC muscular dystrophy, inherited emphysema, familial hypercholesterolemia,  
 CC diabetes, hypopituitarism, adenine deaminase deficiency, alphas-  
 CC antitrypsin deficiency, Gaucher's syndrome, anaemia, infections such as  
 CC HIV infection, high blood pressure, Alzheimer's disease, autoimmune or  
 CC inflammatory disease or graft versus host disease. RV's are capable of  
 CC surviving inactivation in human serum thereby allowing efficient gene  
 CC transfer over prolonged periods of time.

XX Sequence 4832 BP; 1381 A; 1089 C; 1055 G; 1307 T; 0 other;

Query Match 79.0%; Score 790; DB 19; Length 4832;  
 Best Local Similarity 99.4%; Pred. No. 2,6e-236;  
 Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 gatcaagtgatgacgtgtgacacatgatatcaccgacatcaagcaggtgccc 60  
 Db 3656 gatcaagtgatgacgtgtgacacatgatatcaccgacatcaagcaggtgccc 3715  
 QY 61 tcagaagttccacagctccacatctcagttatcatatgatatgctgagga 120  
 Db 3716 tcagaagttccacagctccacatctcagttatcatatgatatgctgagga 3775  
 QY 121 gaagtgcagactatcaggaataatccactggaacttaagtctctcttgcaatg 180  
 Db 3776 gaagtgcagactatcaggaataatccactggaacttaagtctctcttgcaatg 3835  
 QY 3776 gaagtgcagactatcaggaataatccactggaacttaagtctctcttgcaatg 3835  
 QY 181 ggaatcactggtgataaacaacatattttaaccctccaattatgctgatacgcg 240  
 Db 3836 ggaatcactggtgataaacaacatattttaaccctccaattatgctgatacgcg 3895  
 QY 241 ttgaccaccaactatataagatcgcagactcttcgcagtgagttgagtgctgta 3900  
 Db 3896 ttgaccaccaactatataagatcgcagactcttcgcagtgagttgagtgctgta 3955  
 QY 301 tttaaatggtcagacatgacattggaatgagagtaaaacatatacagatcacagat 360  
 Db 3956 tttaaatggtcagacatgacattggaatgagagtaaaacatatacagatcacagat 4015  
 QY 361 tactgtctatcactacttaccatatagtttgacacctgtctccttcaaaagctcga 420  
 Db 4016 tactgtctatcactacttaccatatagtttgacacctgtctccttcaaaagctcga 4075  
 QY 421 tcaactcagaaggaggaatgactgctgagagactcagtggaataatcaaaagtgct 480  
 Db 4076 tcaactcagaaggaggaatgactgctgagagactcagtggaataatcaaaagtgct 4135  
 QY 481 gcaagtgagacttcagaaacaaatgaatgacagagtaactactcagggagtaaatc 540  
 Db 4136 gcaagtgagacttcagaaacaaatgaatgacagagtaactactcagggagtaaatc 4195  
 QY 541 tctgtaccacagcgtatgtagaagaggtctcactcaccagcagtgaggtgagc 600  
 Db 4196 tctgtaccacagcgtatgtagaagaggtctcactcaccagcagtgaggtgagc 4255  
 QY 601 gtgagactctctttttagaatgcaaatgaaggttttcaggggaaatcaagactcct 660  
 Db 4256 gtgagactctctttttagaatgcaaatgaaggttttcaggggaaatcaagactcct 4315

QY 661 cacactgtgtgacatctctctagaccacgctactactcagctcactcgaattcacc 720  
 Db 4316 cacactgtgtgacatctctctagaccacgctactactcagctcactcgaattcacc 4375  
 QY 721 ccagagttggtgcacccagatgccccttagagatgaggttctggctcgagacaga 780  
 Db 4376 ccagagttggtgcacccagatgccccttagagatgaggttctggctcgagacaga 4435  
 QY 781 cctctactgactcagacc 798  
 Db 4436 cctctactgaggtgccc 4453

# RESULT 10

AAV15338 standard; DNA; 4832 BP.

AAV15338;

20-JUL-1998 (first entry)

Human Factor VIII SQN deletion mutant DNA.

Factor VIII; blood clotting; haemophilia A; gene therapy;

retrovirus; vector; human; ss.

Homo sapiens.

Synthetic.

Key Location/Qualifiers

FT CDS 72..445

FT /\*lag= a

PN MO9800542-A2.

PD 08-JAN-1998.

PF 02-JUL-1997; 97WO-US11785.

PR 04-JUN-1997; 97US-0869309.

PR 03-JUL-1996; 96US-0645601.

PR 13-AUG-1996; 96US-0696381.

PA (CHIR) CHIRON CORP.

PI Allen JR, Barber JR, Boder M, Chang SM, Chong K;

PI De La Vega D, Depolo NJ, Greengard J, Hsu DC, Idanez CE;

PI Jolly DJ, Mittelstaedt DM, Prusak CE, Respass JG;

DR WPT: 1998-086967/08.

DR P-PSDB: AAM44372.

PT New replication defective recombinant retroviruses - which express B

PT domain-deleted human factor VIII or human factor IX for the

PS treatment of haemophilia

PS Claim 6; Page 174-175; 236pp; English.

XX This DNA sequence includes a coding region for the B domain  
 CC deletion mutant SQN (see AAM44372) of human factor VIII. The SQN  
 CC mutant is created by fusing Ser-743 to Gln-1638 of native factor  
 CC VIII (see AAM44372) to form a Ser-Gln-Asn (SQN) link between the A2  
 CC and A3 factor VIII domains. When compared to plasmid-derived  
 CC factor VIII, the SQN deletion does not influence the in vivo  
 CC pharmacokinetics, but the reduced size of the molecule appears to  
 CC decrease proteolytic degradation. The invention relates to  
 CC preparations of replication defective recombinant retrovirus (RV)  
 CC expressing a B domain-deleted human factor VIII protein, where the  
 CC recombinant RV is capable of infecting human cells, is resistant  
 CC to degradation by human complement and is capable of inducing  
 CC long-term (at least 30 days and up to 6 months or longer  
 CC post-injection) systemic expression of factor VIII when  
 CC administered to a haemophilia A patient.

XX Sequence 4832 BP; 1381 A; 1089 C; 1055 G; 1307 T; 0 other;  
 SO

Query Match 79.0%; Score 790; DB 19; Length 4832;  
 Best Local Similarity 99.4%; Pred. No. 2.6e+236;  
 Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 gataaaggatgacatctgtgacacatgattatcaccgacatcaagaccaggtgccc 60  
 DB 3656 gatcaaggatgacatctgtgacacatgattatcaccgacatcaagaccaggtgccc 3715  
 QY 61 tcgaagtttccagcctacatctctcagttatcatalcatgatagtctgagggaa 120  
 DB 3716 tcgaagtttccagcctacatctctcagttatcatalcatgatagtctgagggaa 3775  
 QY 121 gaagtgcagactatccaggaatccactggaacttaaggtctcttggcaatgt 180  
 DB 3776 gaagtgcagactatccaggaatccactggaacttaaggtctcttggcaatgt 3835  
 QY 181 ggaatcatctgggataaacaacataattttaacctccaattatgtcgataatcgc 240  
 DB 3836 ggaatcatctgggataaacaacataattttaacctccaattatgtcgataatcgc 3895  
 QY 241 ttggacccaactcatatagatcctgcagcctcttcgacatggatgggctgtga 300  
 DB 3896 ttggacccaactcatatagatcctgcagcctcttcgacatggatgggctgtga 3955  
 QY 301 tttaaatagttgcagatgcatttgggaatgagagtaagaacatataatgcagat 360  
 DB 3956 tttaaatagttgcagatgcatttgggaatgagagtaagaacatataatgcagat 4015  
 QY 361 tactgtatctactcttaccatataatgtttggcaactgtctcttccaagtcgact 420  
 DB 4016 tactgtatctactcttaccatataatgtttggcaactgtctcttccaagtcgact 4075  
 QY 421 tcaactccaaggaggaagtaatgcttggagacccctgaagtataatccaagaatgct 480  
 DB 4076 tcaactccaaggaggaagtaatgcttggagacccctgaagtataatccaagaatgct 4135  
 QY 481 gaaatggaacttccaggaagaacatgaaatgacagagtaactactcaaggagtaaatc 540  
 DB 4136 gaaatggaacttccaggaagaacatgaaatgacagagtaactactcaaggagtaaatc 4195  
 QY 541 tctgttaccagcatgtatgtaagaggttccatctcagacagatcaagatgacatca 600  
 DB 4196 tctgttaccagcatgtatgtaagaggttccatctcagacagatcaagatgacatca 4255  
 QY 601 gtggaactctttttcagaatggaagttttcagaagaaatcaagactcctt 660  
 DB 4256 gtggaactctttttcagaatggaagttttcagaagaaatcaagactcctt 4315  
 QY 661 cacacctgtgtgaactctctagaaccacgctactgactcgtactcgaatcacc 720  
 DB 4316 cacacctgtgtgaactctctagaaccacgctactgactcgtactcgaatcacc 4375  
 QY 721 ccgaagtgtgtgacacagatctgacctgagatgagagttctggtcgagagcaagaa 780  
 DB 4376 ccgaagtgtgtgacacagatctgacctgagatgagagttctggtcgagagcaagaa 4435  
 QY 781 cctctactgactcgagcc 798  
 DB 4436 cctctactgaggtgagcc 4453

RESULT 11

AAT69811 standard; DNA; 5035 BP.

AC AAT69811.

XX 10-AUG-1997 (first entry)

XX

DE Factor VIII-dB695-HCII DNA.

XX Factor VIII-dB695-HCII: heparin cofactor II; blood coagulation;  
 KW blood clotting; procoagulant; anticoagulant; antithrombotic;  
 KW haemophilia; gene therapy; ss.  
 XX

OS Homo sapiens.

XX Key Location/Qualifiers

XX CDS /tag= a

XX misc\_RNA /product= Factor VIII-dB695-HCII

XX /tag= b /product= heparin cofactor II region (aa51-81)

XX WO9718315-A1.

XX 22-MAY-1997.

XX 13-NOV-1996; 96WO-EP04977.

XX 13-NOV-1995; 95US-0558107.

XX (IMMUNO) IMMUNO AG.

XX Voorberg JJ;

XX WPI: 1997-289291/26.

XX P-PSDB: AAM18670.

XX Hybrid Factor VIII with modified activity, comprises region from

XX donor anticoagulant or antithrombotic protein - useful for treatment

XX of coagulation disorders

XX Claim 16; Page 52-60; 96pp; English.

XX A DNA molecule (AAT69811) codes for Factor VIII-dB695-HCII (AAM18670),  
 XX a hybrid protein in which amino acids 712-736 of Factor-dB695  
 XX (Factor VIII del868-1562) B-domain are replaced by amino acids  
 XX 51-80 from the acidic region (and potential thrombin-binding site)  
 XX of human heparin cofactor II (HCII). It was obtd. by PCR  
 XX amplification (see also AAT69812-13) of the HCII acidic region from  
 XX total liver cDNA, fusion to sequences encoding Factor VIII  
 XX aa706-711 and aa737-743, and incorporation of the construct into  
 XX plasmid pCIB-dB695. The hybrid protein, which can be expressed  
 XX using gene therapy techniques, has increased procoagulant activity  
 XX owing to the HCII acidic region, and can be used to treat blood  
 XX coagulation disorders such as haemophilia A.

Sequence 5035 BP; 1484 A; 1127 C; 1110 G; 1314 T; 0 other;

Query Match 79.0%; Score 790; DB 18; Length 5035;  
 Best Local Similarity 99.4%; Pred. No. 2.6e+236;  
 Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 gataaaggatgacatctgtgacacatgattatcaccgacatcaagaccaggtgccc 60  
 DB 4231 gataaaggatgacatctgtgacacatgattatcaccgacatcaagaccaggtgccc 4290  
 QY 61 tcgaagtttccagcctacatctctcagttatcatalcatgatagtctgagggaa 120  
 DB 4291 tcgaagtttccagcctacatctctcagttatcatalcatgatagtctgagggaa 4350  
 QY 121 gaagtgcagactatccaggaatccactggaacttaaggtctcttggcaatgt 180  
 DB 4351 gaagtgcagactatccaggaatccactggaacttaaggtctcttggcaatgt 4410  
 QY 181 ggaatcatctgggataaacaacataattttaacctccaattatgtcgataatcgc 240  
 DB 4411 ggaatcatctgggataaacaacataattttaacctccaattatgtcgataatcgc 4470







Db 6987 ccagagttggtgacacagatgcccctgagatgagagttctggtcgagacacaga 7046  
QY 781 cctctactga 790  
|||||  
Db 7047 cctctactga 7056

## RESULT 14

AAT31031  
ID AAT31031 standard; CDNA: 8967 BP.

AC AAT31031;

DT 19-SEP-1996 (first entry)

DE Factor-VIII full-length cDNA

KW Factor-VIII: retrovirus; vector; haemophilia A; gene therapy; ds.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 110..7165

FT sig\_peptide 110..166

FT mat\_peptide 167..7162

XX W09621035-A2.

XX 11-JUL-1996.

XX 18-DEC-1995; 95WO-US16582.

XX 30-DEC-1994; 94US-0366851.

XX (CHIR) CHIRON VIAGENE INC.

XX Bodner M, Chang S, Chi-Tang Hsu D, De Polo NJ;

XX WPI: 1996-334010/33.

XX P-PSDB: AAM00465.

XX Claim 3; Page 58-68; 86pp; English.

CC A full-length cDNA clone (AAT31031). codes for human Factor-VIII  
(AAM00465), a trace plasma glycoprotein which acts as a cofactor in  
CC conjugation with Factor-IXa in the activation of Factor-X.  
CC Retroviral vectors comprising the full-length cDNA can be efficiently  
CC packaged into infectious retroviral particles. These may be used to  
CC transduce cells either in vivo or ex vivo. Factor-VIII expressed  
CC from such transduced cells will be processed and transported in a  
CC fashion analogous to the expression product of a normal Factor-VIII  
CC gene. Retroviral particles harbouring such vectors will be useful  
CC in the gene therapy of haemophilia A.

XX Sequence 8967 BP; 2841 A; 1898 C; 1833 G; 2395 T; 0 other;

Query Match 79.08; Score 790; DB 17; Length 8967;  
Best Local Similarity 99.48; Pred. No. 3, 6e-236;  
Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 gatcaaggatgacatctgtgacacatgatattacagagatcaagaccaggtgccg 60  
|||||  
Db 6376 gatcaaggatgacatctgtgacacatgatattacagagatcaagaccaggtgccg 6435  
QY 61 tcagaagtctccagcctctacatctcgaattatatacatcgtatagttatgagaa 120  
|||||

Db 6436 tcagaagtctccagcctctacatctcgaattatatacatcgtatagttatgagaa 6495  
QY 121 gaagtgcagacattatcgaggaattccactggaaccttaattgctcttggcaatgt 180  
|||||  
Db 6496 gaagtgcagacattatcgaggaattccactggaaccttaattgctcttggcaatgt 6555  
QY 181 ggattcatctggataaacaacatatattttaaccctccaattatgtctgatacatcgg 240  
|||||  
Db 6556 ggattcatctggataaacaacatatattttaaccctccaattatgtctgatacatcgg 6615  
QY 241 ttgaccaccaactcattatagatcgcagacactcttcgcagtgtgattggtcgtga 300  
|||||  
Db 6616 ttgaccaccaactcattatagatcgcagacactcttcgcagtgtgattggtcgtga 6675  
QY 301 tttaaatagttgcacatgacattgggaatgagagatgaagaacatatcgaatgacagat 360  
|||||  
Db 6676 tttaaatagttgcacatgacattgggaatgagagatgaagaacatatcgaatgacagat 6735  
QY 361 tactgtctatctactacttaaccataatgtttgccaacctggtctccttcaaaagctgact 420  
|||||  
Db 6736 tactgtctatctactacttaaccataatgtttgccaacctggtctccttcaaaagctgact 6795  
QY 421 tcaactccaaggagagagatgaatgctgagagactcaggatgaataatccaagaagtgc 480  
|||||  
Db 6796 tcaactccaaggagagagatgaatgctgagagactcaggatgaataatccaagaagtgc 6855  
QY 481 gcaagtggactcccaagaacatgaagaatgacacagagatgaataatccaagaagtgc 540  
|||||  
Db 6856 gcaagtggactcccaagaacatgaagaatgacacagagatgaataatccaagaagtgc 6915  
QY 541 tctgttaccagatgatgtgaagagttcctcattccacagagatgaataatccaagaagtgc 600  
|||||  
Db 6916 tctgttaccagatgatgtgaagagttcctcattccacagagatgaataatccaagaagtgc 6975  
QY 601 gtgactctctttttcagaatggcaagaatgaagttttcaggggaaatccaagaatcctt 660  
|||||  
Db 6976 gtgactctctttttcagaatggcaagaatgaagttttcaggggaaatccaagaatcctt 7035  
QY 661 cacactgtgtgaactctctagaccaccgttaactgactgctcctctgaattacc 720  
|||||  
Db 7036 cacactgtgtgaactctctagaccaccgttaactgactgctcctctgaattacc 7095  
QY 721 ccagagttggtgacacagatgacctgaggtgtgaggtctggtcgagacacaga 780  
|||||  
Db 7096 ccagagttggtgacacagatgacctgaggtgtgaggtctggtcgagacacaga 7155  
QY 781 cctctactgcagacc 798  
|||||  
Db 7156 cctctactgcagaccgtggtgc 7173

## RESULT 15

AAC87526  
ID AAC87526 standard; DNA: 8967 BP.

AC AAC87526;

DT 13-MAR-2001 (first entry)

DE Human factor VIII cDNA, SEQ ID NO:1.

KW Factor VIII: human; A2 domain; C2 domain; LRP-mediated plasma clearance;

KW receptor-dependent clearance; receptor-independent clearance;

KW haemophilia; half-life; ss.

XX Homo sapiens.

XX W0200071714-A2.

XX 30-NOV-2000.

XX 24-MAY-2000; 2000WO-US14111.

PR 24-MAY-1999: 9905-0135847.  
 XX (AMNA-) AMERICAN NAT RED CROSS.  
 XX  
 PI Saenko EL, Strickland DK.  
 XX  
 DR WPI: 2001-025163/03.  
 DR P-PSDB: AAB48843.

XX Factor VIII mutants having increased half-life useful for treating  
 PT hemophilia, comprise one or more amino acid substitutions in the A2  
 PT and/or C2 domain of factor VIII.

PS Claim 41: Page 86-101; 121pp: English.

XX The invention relates to human factor VIII mutants comprising an amino  
 CC acid substitution at one or more positions in the A2 domain and/or an  
 CC amino acid substitution at one or more positions in the C2 domain.  
 CC The invention also encompasses a factor VIII mutant which lacks a B  
 CC domain (AAB48842). The factor VIII mutants have an increased half-life  
 CC in the bloodstream. The A2 domain mutants exhibit reduced LRP-dependent  
 CC (receptor-dependent) clearance of factor VIII, while C2 domain mutants  
 CC have reduced receptor-independent clearance. The invention also relates  
 CC to a method of using RAP (receptor associated protein), a protein which  
 CC inhibits LRP (low density lipoprotein related protein)-mediated ligand  
 CC internalisation, to increase the half-life of factor VIII. The mutant  
 CC factor VIII proteins, and nucleotides encoding them, are useful  
 CC for treating haemophilia. RAP, LRP-binding RAP mutants or fragments, and  
 CC nucleic acids encoding them may also be used in the treatment of  
 CC haemophilia, in combination with a mutant factor VIII protein or DNA of  
 CC the invention. The invention provides means of increasing the half-life  
 CC of factor VIII by reducing its clearance from plasma. The present  
 CC sequence represents cDNA encoding human factor VIII.

XX Sequence 8967 BP; 2841 A; 1898 C; 1833 G; 2395 T; 0 other;

Query Match 79.0% Score 790: DB 22: Length 8967;  
 Best Local Similarity 99.4% Pred. No. 3.6e-236;

Matches 793: Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 gatcaagtgatgctgttgacacatgatatcaccgacacgaagcaggtgccg 60  
 DB 6376 gatcaagtgatgctgttgacacatgatatcaccgacacgaagcaggtgccg 6435  
 QY 61 tcagaagttcccgccctcaccatctcagttatcaccatgagtcgtgatggaa 120  
 DB 6436 tcagaagttcccgccctcaccatctcagttatcaccatgagtcgtgatggaa 6495  
 QY 121 gaagtcgacactatcgaagaaatccacatgacacttaagtctcttggcaatgt 180  
 DB 6496 gaagtcgacactatcgaagaaatccacatgacacttaagtctcttggcaatgt 6555  
 QY 181 ggaatcacttggaataaacaatattttaaccctccaattatgtcgaataccg 240  
 DB 6556 ggaatcacttggaataaacaatattttaaccctccaattatgtcgaataccg 6615  
 QY 241 ttgacacacacatataagcattgacacacacacacacacacacacacacac 300  
 DB 6616 ttgacacacacatataagcattgacacacacacacacacacacacacacac 6675  
 QY 301 ttaaatagttgcagcattgacacacacacacacacacacacacacacacacac 360  
 DB 6676 ttaaatagttgcagcattgacacacacacacacacacacacacacacacacac 6735  
 QY 361 tactgttcaactcacttaccataatggttgcacactggtctccttcaaaagctcga 420  
 DB 6736 tactgttcaactcacttaccataatggttgcacactggtctccttcaaaagctcga 6795  
 QY 421 tcacctcaagggaggaatgacctggaagacactcaggtgaataatccaaagatgct 480  
 DB 6796 tcacctcaagggaggaatgacctggaagacactcaggtgaataatccaaagatgct 6855

QY 481 gcaagtgacttccagaagaacatgaagtcacaggaagtaactcagggagtaaaac 540  
 DB 6856 gcaagtgacttccagaagaacatgaagtcacaggaagtaactcagggagtaaaac 6915  
 QY 541 tctgtaccagatgatgtgaagagttcctcactccacacacacacacacacacacac 600  
 DB 6916 tctgtaccagatgatgtgaagagttcctcactccacacacacacacacacacac 6975  
 QY 601 gtgactctcttttcagaatgcaagtaagtttccagggagaaatcgaactcctt 660  
 DB 6976 gtgactctcttttcagaatgcaagtaagtttccagggagaaatcgaactcctt 7035  
 QY 661 cacacctgtgtgaactctcttgaccacacacacacacacacacacacacacacac 720  
 DB 7036 cacacctgtgtgaactctcttgaccacacacacacacacacacacacacacacac 7095  
 QY 721 ccagagttggtcac 780  
 DB 7096 ccagagttggtcac 7155  
 QY 781 cctctactgaactcagacc 798  
 DB 7156 cctctactgaaggtggcc 7173

Search completed: January 8, 2002, 17:33:36  
 Job time: 1296 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 8, 2002, 15:13:44 ; Search time 4541.38 Seconds  
(without alignments)  
2366.192 Million cell updates/sec

Title: US-09-740-211-14\_COPY\_4000\_4999

```

Perfect score: 1000
Sequence: 1 gatcaagtgtatctgttg.....cgcagctgtcctgcagagacat 1000

```

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```

Post-processing:  Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

```

Database :

```

1:  em_estfun:*
2:  em_esthbm:*
3:  em_estfn:*
4:  em_estgm:*
5:  em_estpa:*
6:  em_estba:*
7:  em_estov:*
8:  em_estov:*
9:  em_htc:*
10:  gp_est1:*
11:  gp_est2:*
12:  gp_htc:*
13:  gp_gss:*
14:  em_gss_fun1*
15:  em_gss_hum:*
16:  em_gss_inv:*
17:  em_gss_pln:*
18:  em_gss_pro:*
19:  em_gss_rtd:*
20:  em_gss_vrt:*
21:  em_gss_other:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	573.8	57.4	692	11	BG700655	BG700655 602682272
2	506	50.6	720	10	AL601022	AL601022 DKFP313G
3	449.2	44.9	815	10	AI135340	AI135340 AU135340
4	446.4	44.6	943	11	BG1818325	BG1818325 602780212
5	398.8	39.9	482	10	BE650938	BE650938 UI-M-BH3-
6	350.8	35.1	515	10	AV748467	AV748467 AV748467
7	280.4	28.0	639	11	BI067808	BI067808 PGF1h.pK0
8	217.4	21.7	401	11	BE847128	BE847128 UW22B11.y
9	201.4	20.1	444	10	AW489638	AW489638 UI-M-BH3-
10	194	19.4	265	10	AA184901	AA184901 mt93612.r
11	193.4	19.3	384	11	BE852200	BE852200 UW22B11.x
12	193.2	19.3	247	10	AW446518	AW446518 86227 MAR

C	13	190	19.0	431	10	AM456831
C	14	158.8	15.9	927	10	AL552613
C	15	136.6	13.7	745	10	AL577036
C	16	136.2	13.6	965	11	BF302800
C	17	132.6	13.3	616	11	BG641736
C	18	131.8	13.2	864	10	AI334553
C	19	130.6	13.1	942	11	BG175139
C	20	130	13.0	657	10	AI173222
C	21	129	12.9	1038	11	BE911604
C	22	128	12.8	772	11	BI251160
C	23	127.8	12.7	290	11	BE848985
C	24	127	12.6	857	11	BG242396
C	25	125.6	12.5	775	11	BG246010
C	26	125.4	12.5	608	10	BE554133
C	27	121.8	12.2	961	11	BI249890
C	28	121.4	12.1	555	11	BG086480
C	29	121.4	12.1	674	11	BI409154
C	30	120.6	12.1	613	10	AA512671
C	31	119.8	12.0	811	11	BI154893
C	32	119.6	12.0	698	10	AL575850
C	33	118.8	11.9	763	11	BI176281
C	34	117.4	11.7	633	10	AA208846
C	35	116.4	11.6	984	11	BG176124
C	36	115.8	11.6	654	11	BE916565
C	37	115.8	11.6	916	11	BG242699
C	38	115.2	11.5	616	10	BE284409
C	39	114.4	11.4	270	11	BG090843
C	40	114.2	11.4	784	11	BI105565
C	41	112.4	11.2	794	11	BI154670
C	42	112.4	11.2	871	11	BE167957
C	43	111.4	11.1	790	11	BG976390
C	44	110.4	11.0	549	10	BE381295
C	45	110.4	11.0	626	11	BI065410
						AM456831
						UI-M-BH3
						AL552613
						AL577036
						AL577036
						BF302800
						BG641736
						PG1.c.pk0
						AI334553
						BG175139
						60233745
						AI173222
						u63412.x
						BE911604
						602994360
						BI251160
						602994360
						BE848985
						uW01102.y
						602354177
						BG242396
						6023341177
						BG246010
						602338838
						BE554133
						UI38608.y
						602995994
						BI249890
						AI126F11-
						BG086480
						BI409154
						602961479
						AA512671
						u939602.r
						BI154893
						602902882
						AL575850
						AI575850
						BI176281
						602338150
						AA208846
						nmv75a01.r
						BG16124
						602337956
						BE916565
						601667363
						60235528
						BG24269

## ALIGNMENTS

RESULT	1
LOCUS	BG700655
DEFINITION	BG700655 692 bp mRNA EST 07-MAY-2001 602662222f1 NIH_MGC_95 Homo sapiens CDNA clone IMAGE:4815208 5', mRNA sequence.
ACCESSION	BG700655
VERSION	BG700655.1 GI:13970214
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo. 1 (bases 1 to 692)
AUTHORS	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: <a href="mailto:cgabbs@email.nih.gov">cgabbs@email.nih.gov</a> Tissue Procurement: Miklos Palokovits, M.D., Ph.D. CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki Toshiyuki and Piero Carninci (RIKEN) CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Plate: LLAM10711 row: 0 column: 17 High quality sequence stop: 690. Location/Qualifiers
FEATURES	1..692
SOURCE	

## FEATURES

Source



RESULT	3	EST	24-OCT-2000
LOCUS	AU135340	815 bp	MRNA
DEFINITION	AU135340	PLACER Homo sapiens cDNA clone	PLACEL001786 5', MRNA
ACCESSION	AU135340		
VERSION	AU135340.1	GI:10995879	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 815) Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and Iisaga,T.		
TITLE	HRI human cDNA project		
JOURNAL	Unpublished (2000)		
COMMENT	Contact: Takao Iisaga Genomics Laboratory Helix Research Institute 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan Tel: 81-438-52-3851 Fax: 81-438-52-3852 Email: genomics@hri.co.jp HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.		

```

FEATURES
source      Location/Qualifiers
1. 815
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="PLACE1001786"
/clone_1b="PLACE1"
/tissue_t1pe="placenta"
/note="Vector: pME185FL3"
BASE COUNT      200 a      207 c  204 g      197 t
ORIGIN
7. others

```

Query Match	Similarity	Score	DB 10	Length	815
Best Local	454	Conservative	0	Mismatches	8
			0	Indels	0
			0	Gaps	0

```

Qy      581  agcagtcgaatgycgcatcagtcagtcctcttttcgaat 622
      |||||
      512  agcagtcgaatgycgcatcagtcagtcctcttttcgaat 553
Db

```

RESULT	4
BC818325	
LOCUS	943 bp mRNA
DEFINITION	602780212P1 NCI_CGAP_Brn67 Homo sapiens cDNA clone IMAGE:4915634
ACCESSION	BC818325
VERSION	5, mRNA sequence.
KEYWORDS	BC818325.1 GI:14165912
SOURCE	EST.
ORGANISM	human.

ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Homo sapiens	1 (bases 1 to 943)				
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .		National Institutes of Health, Mammalian Gene Collection (MGC)	Unpublished (1999)	
	Contact: Robert Strausberg, Ph.D.				
	Email: <a href="mailto:cgabbs-remail.nih.gov">cgabbs-remail.nih.gov</a>				
	Tissue procurement: David N. Louis, M.D.				
	cDNA library preparation: Life Technologies, Inc.				
	cDNA library Arrayed by: The I.M.A.G.E. Consortium (LNL)				
	DNA Sequencing by: Incyte Genomics, Inc.				
	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:				
	<a href="http://image.llnl.gov">http://image.llnl.gov</a>				
	Plate: LLM10824	row: b	column: 03		
	High quality sequence stop: 629.				

FEATURES	SOURCE	LOCATION/Qualifiers
		1. .943
		/organism="Homo sapiens"
		/db_xref="taxon:9606"
		/clone="IMAGE:4915634"
		/clone_1ib="NCI_CGAP_Brn67"
		/tissue_type="anaplastic oligodendroglioma with 1p/19q loss"
		/lab_host="DH10B (T1 phage-resistant)"
		/note="Organ: brain; Vector: pCMV-SPORT6; Site: 1; NotI: Site_2; Sall: Cloned unidirectionally. Primer: Oligo dT. Average insert size 2.3 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT	252 a	241 c 210 g 240 t
ORIGIN		

Query Match	44.6%	Score 446.4	DB 11	Length 943
Best Local Similarity	99.8%	Pred No. 1	1e-120	
Matches 447	Conservative	0	Mismatches 1	Indels 0
				Gaps 0
Qy	343	aatalcagaatgcacagaattactgattcattccacttaaccataatgtttgcacccgtgic	402	
Db	2	AATATCGATGACACAGATTACTGCTTCATCTCTATTACCAATATATTTGGCCACCTGGTC	61	
Qy	403	tccttcaaaagctgcagcttaccctcccaaggagagatgaatgctcttgagacctcaagltga	462	
Db	62	TCCTTCAAAAGCTGCAGCTTCACTCCCAAGGAGAGAGTAATGCTCGAGACTCAGGTGA	121	
Qy	463	taatccaaaagaatlgctgtcgaagtgaactccagaagaagacaaataaagtcacagaagtaac	522	
Db	122	TAAATCCAAAAGAGGTGTCGCAAGTGAGACTTCGAGAAACATGAAATCCACAGAGTAAAC	181	
Qy	523	tactcagaaggataaactctctgtcttaccagaagatgatgtgaagaggttccatctccag	582	
Db	182	TACTCAGGAGGATAAATATCTCTGCTTACCACAGCATATATGTAAGAGAGTTCCTCATCTCCAG	241	
Qy	583	cagtcagaagatgagcatcagtgagactccctttttcagaatggcacaagttaaagttttca	642	
Db	242	CAGTCAGATGAGCATCAGTGAGACTCTCTTTTTCACAATATGGCAAGTAAAGGTTTTTCA	301	

QY	643	gggaaatcaaaagctctctcaacactggtggtgaactctctctagaccacgcttactgactg	702
DB	302	gggaatcagaagctctctcaacactggtggtgaactctctctagaccacgcttactgactg	361
QY	703	ctaccctcgaattcaacccccagagttgggtgcacgaattgcctctagatgagattct	762
DB	362	ctaccctcgaattcaacccccagagttgggtgcacgaattgcctctagatgagattct	421
QY	763	gggctcggaagcagaggaactctctctga	790
DB	422	gggctcggaagcagaggaactctctctga	449
RESULT	5		
LOCUS	BE650938		
DEFINITION	BE650938	482 bp	mRNA
ACCESSION	BE650938	06-SEP-2000	
VERSION	BE650938		
KEYWORDS	BE650938.1	GI:976762	
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Enkaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus		
AUTHORS	1 (bases 1 to 482)		
TITLE	Bonaldo,M.F., Lennon,G. and Soares,M.B.		
JOURNAL	Normalization and subtraction: two approaches to facilitate gene		
MEDLINE	discovery		
COMMENT	Genome Res. 6 (9), 791-806 (1996)		
	9704477		
	Contact: Chin, H		
	National Institute of Mental Health		
	6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD		
	20892-9643, USA		
	Tel: 301 443 9706		
	Fax: 301 443 9690		
	Email: mstemall.nih.gov		
	cDNA library Preparation: M.B. Soares Lab Clone distribution:		
	Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It		
	should be noted that Bento Soares is generating a small number of		
	additional specialized non-redundant arrays of BMAP cDNAs whose		
	availability will be considered under appropriate and limited		
	collaborative arrangements		
	Seq primer# M13 Reverse.		
	Location/Qualifiers		
	1. 482		
	/organism="Mus musculus"		
	/strain="C57BL/6J"		
	/db_xref="taxon:10090"		
	/clone="U1-M-BH3-asm-g-04-0-U1"		
	/clone_1lb="NIH_BMAP_M_S4"		
	/dev_stage="27-32 days"		
	/lab_host="DH10B (Life Technologies)"		
	/note="Vector: p773D-Pac (Pharmacia) with a modified		
	polylinker. Site_1: Not I; Site_2: Eco RI; The		
	NIH_BMAP_M_S4 library is a subcloned library of a series		
	ultimately derived from a mixture of individually tagged		
	normalized libraries from ten regions of the mouse brain		
	(cerebellum, brain stems, olfactory bulbs, hypothalamus,		
	cortex, amygdala, basal ganglia, pineal gland, striatum,		
	hippocampus) after a series of subtractions to reduce the		
	representation of cDNAs from which ESTs had already been		
	generated. The following serially subtracted libraries		
	were generated in this process: NIH_BMAP_M_S4,		
	NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, NIH_BMAP_M_S3.1,		
	NIH_BMAP_M_S2, NIH_BMAP_M_S1. The subtracted library		
	(NIH_BMAP_M_S4) was constructed as follows: PCR amplified		
	cDNA inserts from NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and		
	NIH_BMAP_M_S3.1 clones from which 3 ESTs had been derivate		
	was used as a driver in a hybridization with a pool of		
	the NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and NIH_BMAP_M_S3.1		

libraries in the form of single-stranded circles. The remaining single-stranded circles (subtracted library) was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (Lifetechnologies) to generate the *NH<sub>4</sub>BMAP<sub>M</sub>-S4* library. This procedure has been previously described (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996) "

Query Match	Best Local Similarity	39.9%	Score 398.8	DB 10	Length 482
Matches 430	Conservative	0	Mismatches 52	Indels 0	Gaps 0
Oy	113	gatgagaagaagtggcagactatcgaagaaatccactggaaccttaatgctctc	172		
Db	1	GATGGGAAGAAGTGGCGTAGATATCAAGGAATTCACACTGAGACCTTAATGGCTTCTTT	60		
Oy	173	ggcaatgtgatattcatcttggatgaataacaaatatttttaacctccaattatgctga	232		
Db	61	GGCAATGTGACACTCATCTGGGATTAAGCATTAATAGTTTAATCTCCAAATATVGGCTCA	120		
Oy	233	tacatccgtttgcaccaactcattatagatcgcagcaactcttcgcatgaggttgatg	292		
Db	121	TATATCCGTTTGGACCCCACTCATCTAGATCCGTAAGTAAGTCTTGCGATGGAGTTGATG	180		
Oy	293	ggcttggatttaaatagttggcagcatgcatcttgaggaaatggagatgaacatatagat	352		
Db	181	GGCTTGGATTTTAACAGATTGCACATACCATVGGATGGAAAGTAAGTAATATACGAT	240		
Oy	353	ggcaagattactgcttcatcacttacttaccaatatgtttgccacctgctccctcaaaa	412		
Db	241	ACACAATACACTCCCTCATCTCTTACACCAATGTGTAAGTCTGCTCTTACCA	300		
Oy	413	ggctcagcttcatcctccaaggaggagtaattgctctggaagacctcaggttgaatacca	472		
Db	301	GGCTCAGCTTCACTCCAGGAAGAGATTAATGCTGGCGACCTCAGTGAATATCCAAA	360		
Oy	473	gaattggcttcaagtggacttccagaagaacttgaagaatcacaggaagtaactactcaga	532		
Db	361	CAATGGTTCAGAGTGGACCTTACAAAAGCAATGAAGTACTCTGGAATTAATACCAAGCA	420		
Oy	533	gtaaaatctctgcttaccagcatgtatgtgaaggagttccatctccagcagtcagaat	592		
Db	421	GTGAATCTCTCTTATACAGCATGTTTGGAAGAAGATTCTTATTTTACAGCAGTCAAGAT	480		
Oy	593	gg 594			
Db	481	GG 482			
RESULT	6				
LOCUS	AV748467				
DEFINITION	AV748467 NPC Homo sapiens cDNA clone NRCAXX05 5', mRNA sequence.	EST	19-OCT-2000		
ACCESSION	AV748467				
VERSION	AV748467.1	GI:10906315			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1 (bases 1 to 515)				
AUTHORS	Song,H., Peng,Y., Gu,Y., Yang,Y., Gao,G., Xiao,H., Xu,X., Li,N., Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Ye,M., Zhang,O., Han,Z., Chen,Z., Hu,R. and Chen,J.				
TITLE	Homo sapiens NPC library cDNA clones				
JOURNAL	Unpublished (2000)				
COMMENT	Contact: Qinghua Zhang				
	Shanghai Institute of Endocrinology, Rui-Jin Hospital				



197 Rui-Ji II Road, Shanghai 200025, P. R. China  
Tel: 86-21-64370045 (ex. 663332)  
Fax: 86-21-64743206  
Email: mbsh@ems.sin.sh.cn  
This clone is available at Shanghai Hematology Institute in  
Shanghai.  
Chinese National Human Genome Center at Shanghai  
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong,  
Shanghai.

```

/note="Vector: pbluescript sk(-):.Site_1: Ecort; Site_2:
XhoI"
BASE COUNT      132 a      114 c      131 g      135 t      3 others
ORIGIN

```

RESULT	7
BI067808	
LOCUS	
DEFINITION	BI067808 639 bp mRNA EST 15-JUN-2001
	pgfin.phk008.h22 normalized chicken fat cdna library gallus gallus
	cdna clone pgfin.phk008.h22 5' similar to gi1667973
	reflmp.032003.11 coagulation factor viii; Factor viii [mus muscul]
	sip1061941FA8 mouse COAGULATION FACTOR VIII PRECURSOR
	(PROCOAGULANT COMPONENT) pir1447004 coagulation factor viii
	precursor - mouse gbl/AA37385.11 (L05573) coagulat1, mRNA sequence
ACCESSION	BI067808
VERSION	BI067808.1 GI:14475330
KEYWORDS	EST.
SOURCE	chicken.
ORGANISM	Gallus gallus
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
	Phasianinae; Gallus.

REFERENCE	1 (bases 1 to 639)
AUTHORS	Cogburn, L. A., Morgan, R. W. and Burnside, J.
TITLE	Chicken ESFs from fat
JOURNAL	Unpublished (2001)
COMMENT	Contact: Larry A. Cogburn University of Delaware

FEATURES	source	Location/Qualifiers
		1. .639
		/organism="Gallus gallus"
		/db_xref="taxon:9031"
		/clone="cgfin.pk008.h22"
		/clone_id="normalized chicken fat cDNA library"
		/sex="Male and Female"
		/tissue_type="fat"
		/lab_host="E.coli EMDH10B"
		/note="Vector: pSPORT1"
BASE COUNT	165 a	190 c 120 t 14 others
ORIGIN		



NIH-BMAP\_M.S4 library. This procedure has been previously described (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)

TAG-LIB=NIH-BMAP\_M.S4  
TAG-TISSUE=brain-stems  
TAG-SEQ=TCATG

BASE COUNT 122 a 80 c 100 g 142 t  
ORIGIN

Query Match 20.1% Score 201.4; DB 10; Length 444;  
Best Local Similarity 82.5%; Pred. No. 1.9e-48;  
Matches 245; Conservative 0; Mismatches 46; Indels 6; Gaps 1;

OY 494 cagaagacatgaagtcacagagtaactactcagagagtaaaactctgtctaccagc 553  
DB 444 CAAAGACATGAAGTACTCGAATAATACCCAGGAGTGAATCTCTTACAGC 385  
OY 554 atglatggaagagttctctcaccagagtcagaatggccatcagtgagctctt 613  
DB 384 ATGTTGTGAAGAGTTCTTATTTCCAGACAGTGAAGTGGCCTGAGTCAAAAT 325  
OY 614 ttccagaatggcaagagtttccaggaatcaaacactctccacacgtggtg 673  
DB 324 TTTATACATGCG-----AAGTTTTTCAGGGGAAATCAGACTCAACACCTATGATG 271  
OY 674 aactcttagaccaccgttactgactgactcgaactcctcaccacagagttggtg 733  
DB 270 AATTTCTAGACCCACCATCTACTCGCTATCTTGAAATTCACCCCGAGTCTGGGAG 211  
OY 734 caccagatggcctgagagtgaggttcgtggtcgagagcagagactcactga 790  
DB 210 CACCAATGCTGTGAGCTTGAGATTCTAGATGTGAGGCCAGCAGCAATCTGA 154

## RESULT 10

AA184901 265 bp mRNA EST 17-FEB-1997  
LOCUS AA184901  
DEFINITION uc09jel2.r1 Soares mouse 3NDMS Mus musculus cDNA clone IMAGE:637486  
5' similar to gi210573 Mus domesticus coagulation factor VIII mRNA  
complete (MUS57), mRNA sequence.

ACCESSION AA184901  
VERSION AA184901.1 GI:1268759  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus.

REFERENCE  
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
Marras, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,  
Geisler, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,  
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,  
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and  
Waterston, R.

TITLE The WashU-HMI Mouse EST Project  
JOURNAL Unpublished (1996)  
COMMENT Contact: Marra M/Mouse EST Project  
WashU-HMI Mouse EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouseest@atson.wustl.edu

High quality sequence stop: 250.  
Seq primer: -28M13 rev2 from Amerisham  
High quality sequence stop: 250.  
Location/Qualifiers  
1. 265

FEATURES  
SOURCE /organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"

/clone="IMAGE:637486"  
/clone\_lib="Soares mouse 3NDMS"  
/sex="male"  
/tissue\_type="Spleen"  
/dev\_stage="4 weeks"  
/lab\_host="DH10B"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified  
polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA  
was primed with a Not I - oligo(dT) primer [5',  
TGTTACCAATCTGAAGTGGAGGCGCCGCTGTTTTTTTTTTTTTTTTTT  
3']; double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not I  
and Eco RI sites of the modified pT73 vector. RNA  
provided by Dr. Bertrand Jordan. Library went through  
three rounds of normalization, and was constructed by  
Bento Soares and M. Fatima Bonaldo."

BASE COUNT 82 a 63 c 53 g 67 t  
ORIGIN

Query Match 19.4% Score 194; DB 10; Length 265;  
Best Local Similarity 86.3%; Pred. No. 2.5e-46;  
Matches 226; Conservative 0; Mismatches 35; Indels 1; Gaps 1;

OY 375 acttaccatatttggccactgctccttcaaaagctgactcactccaagga 434  
DB 1 ACTTCACCAATGTTTGTCTTACTGTTGCTCCTTCACAAAGCTGCACTTCACCTCAGGAA 60  
OY 435 ggaatgatgctgagagacttaagtgaataatccaaagtgctcagtgagcttc 494  
DB 61 GGACTAATGCTTGCGGACCTCAGTGAATGATCCAAACAAATGTTGCAAGT-GACTTAC 119  
OY 495 agaagacaatgaagtcacagagtaactcagaggaagtaaaactctctgcttaccaga 554  
DB 120 AAAAGACATGAATGCACTCGAATAATTAACCCAGGAGTGAATTCCTCTTACAGACA 179  
OY 555 tgaatggaagagttctcctcctcaccagagtcagaatgagccatcagtgagctctt 614  
DB 180 TGTGTTGAAGAGTGTCTTATTTCCAGCAGTCAAGATGAGGCAATCAGCTCAAAATTT 239  
OY 615 ttcaagaatggcaagagtaaggt 636  
DB 240 TATACATGCGCAAGGTAAAGGT 261

## RESULT 11

BE852200 384 bp mRNA EST 26-SEP-2000  
LOCUS BE852200  
DEFINITION uc22b11.x1 Soares mouse 3NDMS Mus musculus cDNA clone IMAGE:3417405  
3' similar to SW:FBX\_MOUSE Q06194 COAGULATION FACTOR VIII PRECURSOR  
; mRNA sequence.

ACCESSION BE852200  
VERSION BE852200.1 GI:10310464  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus.

REFERENCE  
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)

JOURNAL Other ESTs: uc22b11.y1  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaaps-r@mail.nih.gov  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.lnl.gov) for further information.  
MGI:1093217.  
Location/Qualifiers  
1. 384

FEATURES  
SOURCE /organism="Mus musculus"  
/strain="C57BL/6J"







GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 8, 2002, 17:23:52 : Search time 177.22 Seconds

(without alignments)  
1277.948 Million cell updates/sec

Title: US-09-740-211-14\_COPY\_4000\_4999

Perfect score: 1000

Sequence: 1 gatcaagtgatcgtctgtg.....cgcagctgcctgcagacat 1000

Scoring table: IDENTITY\_NUC

Gapop 10.0, Gapext 1.0

Searched: 351203 seqs, 11323899 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents\_NA:\*

1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS.COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1000	100.0	4999	4	US-09-470-618-14
2	1000	100.0	4999	4	US-09-364-862-14
3	954.4	95.4	11933	4	US-09-470-618-13
4	954.4	95.4	11933	4	US-09-364-862-13
5	790	79.0	4629	2	US-08-484-891-7
6	790	79.0	4670	3	US-08-717-294-41
7	790	79.0	5035	2	US-08-882-083-1
8	790	79.0	5036	2	US-08-558-107-1
9	790	79.0	5035	3	US-09-243-539-1
10	790	79.0	6999	1	US-08-276-594A-1
11	790	79.0	7056	1	US-08-121-202-1
12	790	79.0	8967	1	US-08-366-851A-1
13	790	79.0	9009	1	US-07-864-004B-3
14	790	79.0	9009	1	US-08-251-937A-3
15	790	79.0	9009	1	US-08-212-133A-1
16	790	79.0	9009	1	US-08-474-503-1
17	790	79.0	9009	2	US-08-670-707A-1
18	790	79.0	9009	4	US-09-037-601-1
19	790	79.0	9009	5	PCT-US93-03275-3
20	790	79.0	9009	5	PCT-US94-13200-1
21	790	79.0	9354	1	US-08-683-839B-2
22	788.4	78.8	8241	6	5171844-1
23	634.8	63.5	7493	1	US-08-212-133A-7
24	634.8	63.5	7493	1	US-08-474-503-5
25	634.8	63.5	7493	2	US-08-670-707A-5
26	634.8	63.5	7493	4	US-09-037-601-5
27	634.8	63.5	7493	5	PCT-US94-13200-5

28	614	61.4	7032	4	US-09-324-867-1	Sequence 1, Appl
29	594.8	59.5	4334	2	US-08-670-707A-38	Sequence 38, Appl
30	594.8	59.5	4334	4	US-09-037-601-38	Sequence 38, Appl
31	594.8	59.5	6402	2	US-08-670-707A-36	Sequence 36, Appl
32	594.8	59.5	6402	4	US-09-037-601-36	Sequence 36, Appl
33	507.6	50.8	4451	3	US-08-717-294-42	Sequence 42, Appl
34	164.4	16.4	6909	2	US-08-804-196-1	Sequence 1, Appl
35	164.4	16.4	6909	2	US-08-658-340-1	Sequence 1, Appl
36	159.6	16.0	6585	3	US-08-746-111-26	Sequence 26, Appl
37	151.4	15.1	4999	4	US-09-470-618-14	Sequence 4, Appl
38	151.4	15.1	4999	4	US-09-364-862-14	Sequence 14, Appl
39	142.6	14.3	11933	4	US-09-470-618-13	Sequence 13, Appl
40	142.6	14.3	11933	4	US-09-364-862-13	Sequence 13, Appl
41	135.6	13.6	8299	1	US-08-462-014-2	Sequence 2, Appl
42	135.6	13.6	8299	3	US-08-923-137-3	Sequence 3, Appl
43	135.6	13.6	8299	4	US-08-973-334-5	Sequence 5, Appl
44	135.6	13.6	8299	4	US-09-563-869A-5	Sequence 5, Appl
45	135.6	13.6	8299	4	US-09-563-869A-5	Sequence 5, Appl

#### ALIGNMENTS

```
RESULT 1
US-09-470-618-14
: Sequence 14, Application US/09470618
: Patent No. 6200560
: GENERAL INFORMATION:
: APPLICANT: Coulo, Linda B.
: APPLICANT: Colosi, Peter C.
: TITLE OF INVENTION: Adeno-Associated Vectors for Expression of Factor VIII
: FILE REFERENCE: Avigen-04082
: CURRENT APPLICATION NUMBER: US/09/470, 618
: CURRENT FILING DATE: 1999-12-22
: EARLIER APPLICATION NUMBER: 09/364, 862
: EARLIER FILING DATE: 1999-07-30
: EARLIER APPLICATION NUMBER: 60/125, 974
: EARLIER FILING DATE: 1999-03-24
: EARLIER APPLICATION NUMBER: 60/104, 994
: EARLIER FILING DATE: 1998-10-20
: NUMBER OF SEQ ID NOS: 15
: SOFTWARE: Patentln Ver. 2.0
: SEQ ID NO 14
: LENGTH: 4999
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-470-618-14

Query Match      100.0%; Score 1000; DB 4; Length 4999;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gatcaagtgatcgtctgtgacacatgatatacagcgatcaagaccagggtgcgcg 60
   |||||
DB 4000 gatcaagtgatcgtctgtgacacatgatatacagcgatcaagaccagggtgcgcg 4059

QY 61 tcagaagttccagcctctacatctcagttatcatcatatagatctgtggaa 120
   |||||
DB 4060 tcagaagttccagcctctacatctcagttatcatcatatagatctgtggaa 4119

QY 121 gaagtcgagcctctacaggaatctcacgcggaacctatgtctcttggcaatgt 180
   |||||
DB 4120 gaagtcgagcctctacaggaatctcacgcggaacctatgtctcttggcaatgt 4179

QY 181 gattcatctggataaacaacaatattttaaaccctcaatttgcgatacatcgcg 240
   |||||
DB 4180 gattcatctggataaacaacaatattttaaaccctcaatttgcgatacatcgcg 4239

QY 241 ttgcaccaactcatatagatctgcagacactcttcgatggatgtgagctgtgta 300
```

```

|||||
Db 4240 ttgacccaactatattatgaatcgaacaccccttcgcatgagttgagctgtga 4299
QY 301 tttaaatgttgacagcatgcatctgggaatgagagaaacaaatcacagatgacagat 360
Db 4300 tttaaatgttgacagcatgcatctgggaatgagagaaacaaatcacagatgacagat 4359
QY 361 tactgtctacatctactattacaaatattgttgcacactggtctctctcaaaagtcact 420
Db 4360 tactgtctacatctactattacaaatattgttgcacactggtctctctcaaaagtcact 4419
QY 421 tcacctccaagggagagtaatgctctggagacccctcaggtgaataatccaaaagatggt 480
Db 4420 tcacctccaagggagagtaatgctctggagacccctcaggtgaataatccaaaagatggt 4479
QY 481 gcaagttgacttccagagacaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatg 540
Db 4480 gcaagttgacttccagagacaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatg 4539
QY 541 tctgtctacagcatgtatgtgaaagagttctcactcaccagatcagaatgagcatca 600
Db 4540 tctgtctacagcatgtatgtgaaagagttctcactcaccagatcagaatgagcatca 4599
QY 601 gtgagactctcttttcaagaatgagaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatg 660
Db 4600 gtgagactctcttttcaagaatgagaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatg 4659
QY 661 caacccctgtgtgacacactctcagacccacccgttactcactcactcactcactcactcactcactc 720
Db 4660 caacccctgtgtgacacactctcagacccacccgttactcactcactcactcactcactcactcactc 4719
QY 721 ccagagtttgagtcacagaaatgtccctgagagatgaggtctctggtcgaagacagga 760
Db 4720 ccagagtttgagtcacagaaatgtccctgagagatgaggtctctggtcgaagacagga 4779
QY 781 cctctactgactcgaagcctaataagaaattatttctctgcaatagtggtttggtt 840
Db 4780 cctctactgactcgaagcctaataagaaattatttctctgcaatagtggtttggtt 4839
QY 841 ttgtgtgagccgacgaagaaacccctagatgagttgagttgagttgagttgagttgagttgagttg 900
Db 4840 ttgtgtgagccgacgaagaaacccctagatgagttgagttgagttgagttgagttgagttgagttg 4959
QY 901 gctgcgtcactgagccggcgacgaagaaatgtgcgcgaagccggcgcttgcgcggcg 960
Db 4900 gctgcgtcactgagccggcgacgaagaaatgtgcgcgaagccggcgcttgcgcggcg 4959
QY 961 cctcagtgagcgaagcgaagccggcgacgtgcgtcgaagacat 1000
Db 4960 cctcagtgagcgaagcgaagccggcgacgtgcgtcgaagacat 4999

```

```

; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-364-862-14

Query Match
Best local similarity 100.0%; Score 1000; DB 4; Length 4999;
Matches 1000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gatcaagtgtagatctgttgcacccaatgatatcaccgcatcacaagaccaggtgccc 60
Db 4000 gatcaagtgtagatctgttgcacccaatgatatcaccgcatcacaagaccaggtgccc 4059
QY 61 tcagaaagttctcagcctctacatctctcagttatcaatcagatagttcttgaggaa 120
Db 4060 tcagaaagttctcagcctctacatctctcagttatcaatcagatagttcttgaggaa 4119
QY 121 gaagtgcagacttaccgaggaatccactggaaccttaatggtctctcttgcaatgt 180
Db 4120 gaagtgcagacttaccgaggaatccactggaaccttaatggtctctcttgcaatgt 4179
QY 181 ggaatcattgggataaacaacataattttaacctcaattatgtcgtatataccg 240
Db 4180 ggaatcattgggataaacaacataattttaacctcaattatgtcgtatataccg 4239
QY 241 ttggacccaactcatatagatctcgcagacactcttcgcagaggttgatggtgtga 300
Db 4240 ttggacccaactcatatagatctcgcagacactcttcgcagaggttgatggtgtga 4299
QY 301 tttaaatagttgacagatgccattgggaaatgagaaatgaagacaatatacagatgacagat 360
Db 4300 tttaaatagttgacagatgccattgggaaatgagaaatgaagacaatatacagatgacagat 4359
QY 361 tactgtctacatctactattacaaatattgttgcacactggtctctctcaaaagtcact 420
Db 4360 tactgtctacatctactattacaaatattgttgcacactggtctctctcaaaagtcact 4419
QY 421 tcacccccaagggagagaaatgagaaatgagaaatgagaaatgagaaatgagaaatgagaaatgagaaatg 480
Db 4420 tcacccccaagggagagaaatgagaaatgagaaatgagaaatgagaaatgagaaatgagaaatgagaaatg 4959
QY 481 gcaagtggaacttccagaaacaaatgaaatgcaagagtaactactcagggagtaaaatc 540
Db 4480 gcaagtggaacttccagaaacaaatgaaatgcaagagtaactactcagggagtaaaatc 4539
QY 541 tctgtctacacagcatgtatgtgaaagagttcctcactcaccagcagtcagaatgacatca 600
Db 4540 tctgtctacacagcatgtatgtgaaagagttcctcactcaccagcagtcagaatgacatca 4599
QY 601 gtgagactctcttttcaagaatgagaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatg 660
Db 4600 gtgagactctcttttcaagaatgagaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatg 4659
QY 661 caacactgtgtgaaactcctcagacccacgttactgactcgtactcgtactcgtactcgtactcgtactcgtactc 720
Db 4660 caacactgtgtgaaactcctcagacccacgttactgactcgtactcgtactcgtactcgtactcgtactcgtactc 4719
QY 721 ccagagtttgagtcacagaaatgtccctgagagatgaggtctctggtcgaagacagga 780
Db 4720 ccagagtttgagtcacagaaatgtccctgagagatgaggtctctggtcgaagacagga 4779
QY 781 cctctactgactcgaagcctaataagaaattatttctctgcaatagtggtttggtt 840
Db 4780 cctctactgactcgaagcctaataagaaattatttctctgcaatagtggtttggtt 4839
QY 841 ttgtgtgagccgacgaagaaacccctagatgagttgagttgagttgagttgagttgagttgagttgagttg 900
Db 4840 ttgtgtgagccgacgaagaaacccctagatgagttgagttgagttgagttgagttgagttgagttgagttg 4959
QY 901 gctgcgtcactgagccggcgacgaagaaatgtgcgcgaagccggcgcttgcgcggcg 960

```



Db 4900 gctcgtcctcctgagcgccgagcaccagaagctccgagccggccttcccgccg 4955  
Oy 961 cctcagtgagcgagcgagcgccgagctgctcagagacat 1000  
Db 4960 cctcagtgagcgagcgagcgccgagctgctcagagacat 4999

## RESULT 3

US-09-470-618-13

Sequence 13, Application US/09470618

Patent No. 6200560

GENERAL INFORMATION:

APPLICANT: Couto, Linda B.

APPLICANT: Colosi, Peter C.

TITLE OF INVENTION: Adeno-Associated Vectors for Expression of Factor VIII

FILE REFERENCE: Avigen-04082

CURRENT APPLICATION NUMBER: US/09/470,618

CURRENT FILING DATE: 1999-12-22

EARLIER APPLICATION NUMBER: 09/364,862

EARLIER FILING DATE: 1999-07-30

EARLIER APPLICATION NUMBER: 60/125,974

EARLIER FILING DATE: 1999-03-24

EARLIER APPLICATION NUMBER: 60/104,994

EARLIER FILING DATE: 1998-10-20

NUMBER OF SEQ ID NOS: 15

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 13

LENGTH: 11933

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: Synthetic

US-09-470-618-13

Query Match 95.4%; Score 954.4; DB 4; Length 11933;  
Best Local Similarity 97.9%; Pred. No. 4.3e-310;  
Matches 979; Conservative 0; Mismatches 16; Indels 5; Gaps 1;

Oy 1 gataaagtgatctgttggcacaatgatatcaccgacatcaagcagggcgccg 60  
Db 3975 gataaagtgatctgttggcacaatgatatcaccgacatcaagcagggcgccg 4034  
Oy 61 tcagaagttccagcctcaatctcctgattatcatatgattatgattgagga 120  
Db 4035 tcagaagttccagcctcaatctcctgattatcatatgattatgattgagga 4094  
Oy 121 gaagtgacagactatcgagaatccactcggaacacttaattgtctcttggcaatgt 180  
Db 4095 gaagtgacagactatcgagaatccactcggaacacttaattgtctcttggcaatgt 4154  
Oy 181 ggaattcctggggaagaacaacatctttaaaccctcaattatgtctgataacccg 240  
Db 4155 ggaattcctggggaagaacaacatctttaaaccctcaattatgtctgataacccg 4214  
Oy 241 tttagcaccacatcatatagatctgacagacactcttcgatgagtgatggcgctga 300  
Db 4215 tttagcaccacatcatatagatctgacagacactcttcgatgagtgatggcgctga 4274  
Oy 301 tttaaatagtgcagacatgacatctggaatgagagtaagaacaatatcagatgcacagat 360  
Db 4275 tttaaatagtgcagacatgacatctggaatgagagtaagaacaatatcagatgcacagat 4334  
Oy 361 tactgtctcaatcctattacacatattgtttgcaactgtgtctcctcaaaagctgcact 420  
Db 4335 tactgtctcaatcctattacacatattgtttgcaactgtgtctcctcaaaagctgcact 4394  
Oy 421 tcacctccaagggaggaatgctcctgagagcctcaggtgaataatccaaagatggct 480  
Db 4395 tcacctccaagggaggaatgctcctgagagcctcaggtgaataatccaaagatggct 4454  
Oy 481 gcaagtgagcttcagagaacatgaaagttcagagagtaactactcagggagtaaaatc 540

Db 4455 gcaagtgagcttcagagaacatgaaagttcagagagtaactactcagggagtaaaatc 4514  
Oy 541 tctgttaccagacatgatatgaaaggatctcctcaatccagcagttcaagatggcaca 600  
Db 4515 tctgttaccagacatgatatgaaaggatctcctcaatccagcagttcaagatggcaca 4574  
Oy 601 gtgacatctcttttcaagaatgaggaatgattttcagggaatcaaacctcct 660  
Db 4575 gtgacatctcttttcaagaatgaggaatgattttcagggaatcaaacctcct 4634  
Oy 661 cacaacgtgtgtaactctctagacccacgcttactgactgctacccctgaattcacc 720  
Db 4635 cacaacgtgtgtaactctctagacccacgcttactgactgctacccctgaattcacc 4694  
Oy 721 ccagagttgggtgacacagatgccctcgagagatgaggtttggtcgcgagagcagga 780  
Db 4695 ccagagttgggtgacacagatgccctcgagagatgaggtttggtcgcgagagcagga 4754  
Oy 781 cctcactgactgagccttaataaggaatttatcttcatgtaatgattgtgtt 840  
Db 4755 cctcactgactgagccttaataaggaatttatcttcatgtaatgattgtgtt 4809  
Oy 841 ttgtgtgctggcggcaggaacccctagtgaggttggcactcctctcgcgcgtc 900  
Db 4810 ttgtgtgctggcggcaggaacccctagtgaggttggcactcctctcgcgcgtc 4869  
Oy 901 gctcgtcactgagcgccggcgacccaaggtcgccggcggcggttggccggcgcg 960  
Db 4870 gctcgtcactgagcgccggcgacccaaggtcgccggcggcggttggccggcgcg 4929  
Oy 961 cctcagtgagcgagcgagcgccagctgcttcagagacat 1000  
Db 4930 cctcagtgagcgagcgagcgccagctgcttcagagacat 4969

## RESULT 4

US-09-364-862-13

Sequence 13, Application US/09364862

Patent No. 6221349

GENERAL INFORMATION:

APPLICANT: Couto, Linda B.

APPLICANT: Colosi, Peter C.

TITLE OF INVENTION: ADENO-ASSOCIATED VECTORS FOR EXPRESSION OF FACTOR VIII

FILE REFERENCE: AVIGEN-03743

CURRENT APPLICATION NUMBER: US/09/364,862

CURRENT FILING DATE: 1999-07-30

EARLIER APPLICATION NUMBER: 60/125,974

EARLIER FILING DATE: 1999-03-24

EARLIER APPLICATION NUMBER: 60/104,994

EARLIER FILING DATE: 1998-10-20

NUMBER OF SEQ ID NOS: 14

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 13

LENGTH: 11933

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: synthetic

US-09-364-862-13

Query Match 95.4%; Score 954.4; DB 4; Length 11933;  
Best Local Similarity 97.9%; Pred. No. 4.3e-310;  
Matches 979; Conservative 0; Mismatches 16; Indels 5; Gaps 1;

Oy 1 gataaagtgatctgttggcacaatgatatcaccgacatcaagcagggcgccg 60  
Db 3975 gataaagtgatctgttggcacaatgatatcaccgacatcaagcagggcgccg 4034  
Oy 61 tcagaagttccagcctcaatctcctgattatcatatgattatgattgagga 120

```

Db 4035 tcaagaattccagcctctacatctcagttatcatcatgatatgctgagaa 4094
QY 121 gaattggcagactatcagaataatccactggagacttaagtctcttggcaatgt 180
Db 4095 gaattggcagactatcagaataatccactggagacttaagtctcttggcaatgt 4154
QY 181 ggaattcctgggataaacaacataatcttaaccctcccaatattatgctgatacag 240
Db 4155 ggaattcctgggataaacaacataatcttaaccctcccaatattatgctgatacag 4214
QY 241 ttggaccacactatataatgcatctgcacactcttcgcataagttgagtgctgta 300
Db 4215 ttggaccacactatataatgcatctgcacactcttcgcataagttgagtgctgta 4274
QY 301 tttaaatagttgcagcatgcatctggaatggaagtaagaacatcaatcaagatcagat 360
Db 4275 tttaaatagttgcagcatgcatctggaatggaagtaagaacatcaatcaagatcagat 4334
QY 361 tactgctcatcctactcttaacataatggttgcacactgctcctctcaaaagctgact 420
Db 4335 tactgctcatcctactcttaacataatggttgcacactgctcctctcaaaagctgact 4394
QY 421 tcacctcgaaggagagtaatcctcctggagacactcaggtgaataatccaaagaagtgct 480
Db 4395 tcacctcgaaggagagtaatcctcctggagacactcaggtgaataatccaaagaagtgct 4454
QY 481 gcaagtgagacttcagaagaacaaatgaagtacagagtaactactcagaagtaaaac 540
Db 4455 gcaagtgagacttcagaagaacaaatgaagtacagagtaactactcagaagtaaaac 4514
QY 541 tctgcttacagagatgatatgtagaagatctcccatctccagatcagaatgagcatca 600
Db 4515 tctgcttacagagatgatatgtagaagatctcccatctccagatcagaatgagcatca 4574
QY 601 gtgagctctcttttcaagaatgcaaaagtgttttcaaggaaatacaagactcct 660
Db 4575 gtgagctctcttttcaagaatgcaaaagtgttttcaaggaaatacaagactcct 4634
QY 661 cacactgtgtgaactctctcagaaccacggtactgactgctactcgaatcacc 720
Db 4635 cacactgtgtgaactctctcagaaccacggtactgactgctactcgaatcacc 4694
QY 721 caagaattgggtgcacacattgccttgagatgagtgctcgtgagtgagacagaa 780
Db 4695 caagaattgggtgcacacattgccttgagatgagtgctcgtgagtgagacagaa 4754
QY 781 cctactactgactcgaataaagaataattatctcatgcaatgagtgtgtgtt 840
Db 4755 cctactactgactcgaataaagaataattatctcatgcaatgagtgtgtgtt 4809
QY 841 ttgtgtgtgcgcgcaggaacccctagtgatgagtggtgcacactcctctgcgcgtc 900
Db 4810 ttgtgtgtgcgcgcaggaacccctagtgatgagtggtgcacactcctctgcgcgtc 4869
QY 901 gctcgtcactgagcgcgcgcgcacaaagtgccgcagacccgcggttttgcgcgcgc 960
Db 4870 gctcgtcactgagcgcgcgcgcgcacaaagtgccgcagacccgcggttttgcgcgcgc 4929
QY 961 cctcagtgagcagcagcgcgcgcagctgctcgcaggaat 1000
Db 4930 cctcagtgagcagcagcgcgcgcagctgctcgcaggaat 4969

```

RESULT 5  
 US-08-484-891-7  
 Sequence 7, Application US/08484891  
 Patent No. 5935935

GENERAL INFORMATION:  
 APPLICANT: Connelly, Sheila  
 APPLICANT: Kalesko, Michael  
 APPLICANT: Smith, Theodore  
 TITLE OF INVENTION: Adenoviral Vectors for

```

? TITLE OF INVENTION: Treatment of Hemophilia
? NUMBER OF SEQUENCES: 7
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Carella, Byrne, Bain, Gillilan,
? ADDRESSEE: Cecchi, Stewart & Olstein
? STREET: 6 Becker Farm Road
? CITY: Roseland
? STATE: New Jersey
? COUNTRY: USA
? ZIP: 07068
? COMPUTER READABLE FORM:
? MEDIUM TYPE: 3.5 inch diskette
? COMPUTER: IBM PS/2
? OPERATING SYSTEM: MS-DOS
? SOFTWARE: WordPerfect 5.1
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/484,891
? FILING DATE: 07-JUN-1995
? CLASSIFICATION: 514
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/218,335
? FILING DATE: 25-MAR-1994
? APPLICATION NUMBER: 08/074,920
? FILING DATE: 10-JUN-1993
? ATTORNEY/AGENT INFORMATION:
? NAME: Olstein, Elliot M.
? REGISTRATION NUMBER: 24,025
? REFERENCE/DOCKET NUMBER: 271010-273
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 201-994-1700
? TELEFAX: 201-994-1744
? INFORMATION FOR SEQ ID NO: 7:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 4629 bases
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: CDNA primer
? FEATURE:
? NAME/KEY: Factor VIII CDNA with
? NAME/KEY: B domain deleted
? US-08-484-891-7

Query Match 79.0%; Score 790; DB 2; Length 4629;
Best Local Similarity 99.4%; Pred. No. 4.4e-255;
Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 gataaagttgatatgttggcacaatgatatctacgagcaagaaccaggtgccc 60
Db 3627 GATCAAGGTGATCTGTTGGCACCACCAATGATTATTCACGGCATCAAGACCCAGGTCGCCG 3686
QY 61 tcaagaattccagcctctacatctcagttatcatcatgatatgtagtcttgaaggaa 120
Db 3687 TCAGAAGTTCACAGCCTCTACATCTCTCAGTTATATCATGTATGATGCTTGATGGGAA 3746
QY 121 gaattggcagactatcagaagaatccactggagacttaagtctcttggcaatgt 180
Db 3747 GAAGTGGCAGACTTATCGAGGAATTCACACTGGAACCTTAATGCTCTTTGGCAATGT 3806
QY 181 ggaattcctgggataaacaacataatcttaaccctcccaatattatgctgatacag 240
Db 3807 GGATTCAATCTGGATTAACAACAAATATTTTAACCCCTCCAAATTAATGCTTCGATACATCCG 3866
QY 241 ttggaccacactatataatgcatctgcacactcttcgcataagttgagtgctgta 300
Db 3867 TTTCACACCAACTATTAATAGCATTTCCAGCACTCTTTCGATGAGTGTGATGGCTGTGA 3926
QY 301 tttaaatagttgcagcatgcatctggaatggaagtaagaacataatcagaatcagat 360
Db 3927 TTTAAATAGTTGCAGCATGCGCATGGAATGGAAGTAAGAAAGCAATATCAATGCACAGAT 3986
QY 361 tactgctcatcctactcttaacataatggttgcacactgctcctctcaaaagctgact 420

```

Db	3987	TACGCTTCACTCTACTTTCACCAATATGTTTGGCACCTGGTCTCTTTCAAAAGCTGCAC	4046
OY	421	tcacctccaaggaagagatcatcctcgagagacctcaagtgaaataccaaaagatggct	480
Db	4047	TCACCTCCAGGAGGAGATATGCTCGGAGACCTCAGGTAATATCCAAAAGATGGCT	4106
OY	481	gcaagtgagactccagaagacaatagaagtcaacaggatgaactactccaggaagfzaaatt	540
Db	4107	GCAAGTGGACTCTCCAGAAAGCAATGAAAGTACACAGGATTAATCTCCAGGAGATAAATTC	4166
OY	541	tcctgtaccagaagatgctgaaagagatcccatctccatccaagatccaagatgccccatca	600
Db	4167	TCTGCTTACACGACATGATATGAAAGAGTTTCTCATCTCCACAGTCAAGATGGGCATCA	4226
OY	601	gtggacatctcttttccaagatggcaaatgaagtgtttccaaggaaatcaaacactct	660
Db	4227	GTGACATCTCTTTTTCAGAAATGGCAAAAGTAAAGTTTTCAGGAAATCAAGACTCCTT	4286
OY	661	caacactgtgtgtgaactctcttagaaccacccgcttaactgaactctgaattcaccc	720
Db	4287	CACACCTGTGTGAACCTCTATAGACCCACCGTATACAGCTCGGTACTCTTCGAAATTCACC	4346
OY	721	ccagaagttgggttcgacccaagattggccctggaagatggagagttcttgagcttgcgagcaaga	780
Db	4347	CCAGAGTGGGGTGACACCAAGATGGCCCTGACAGATGAGAGTTTGGGCTGGCAGGACACAGA	4406
OY	781	ccttaactgaactcgagcc 798	
Db	4407	CCTCTACTGAGGGTGGCC 4424	

RESULT 6  
 US-08-717-294-41  
 Sequence 41, Application US/08717294  
 Patent No. 6114148  
 GENERAL INFORMATION:  
 APPLICANT: SEED, BRIAN  
 APPLICANT: HAAS, JOEGEN  
 TITLE OF INVENTION: HIGH LEVEL EXPRESSION OF  
 TITLE OF INVENTION: PROTEINS  
 NUMBER OF SEQUENCES: 110  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Clark & Elbing, LLP  
 STREET: 176 Federal Street  
 CITY: Boston  
 STATE: MA  
 COUNTRY: USA  
 ZIP: 02110  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FASTSD for Windows Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/717,294  
 FILING DATE: 20-SEP-1996  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER:  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Elbing, Karen L  
 REGISTRATION NUMBER: 35,238  
 REFERENCE/DOCKET NUMBER: 00786/345001  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 617-428-7020  
 TELEFAX: 617-428-7045  
 TELEX:  
 INFORMATION FOR SEQ ID NO: 41:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 4670 base pairs  
 TYPE: nucleic acid

```

; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
US-08-717-294-41

```

Query Match	79.0%;	Score 790;	DB 3;	Length 4670;
Best Local Similarity	99.4%;	Pred. No. 4.4e-255;		
Matches 793;	Conservative	0;	Mismatches 5;	Indels 0;
				Gaps 0.

Qy	1	gatacaaggtagcatcgttgggaccaaigtatataccaaggcatcaagaaccagagtgccg	60
Db	3653	gattcaagatgagatctgttgacacacatgattttacaggacatcaagaccaggatgccc	3712
Qy	61	tcaagaattccagaccttaacatccacagttatacatatgatatgtcttgaibggaa	120
Db	3713	ttcagaagttccacagccttactatctctcagtttatcatatgatatgtcttgaibggaa	3772
Qy	121	gaagtgcaagactatcgaggaatctcaatgaaaccttaatgtctctttggcaat	180
Db	3773	gaagtgcaagactatcgaggaatctcaatgaaaccttaatgtctctttggcaatgt	3832
Qy	181	ggatcatctcgggaataaacaacatcttttaacctccaattctgcgtacataccg	240
Db	3833	ggatcatctcgggaataaacaacatcttttttaacctccaattttgcgtatracattccg	3892
Qy	241	tttgaccacaactatataatgcatctgcagacactcttcgcatlbgagltgaibggcgtga	300
Db	3893	tttgaccaccactatattatgatgtgcagacactcttcgcatlbgagltgaibggcgtga	3952
Qy	301	tttaaatagltgtagcaatgcacatttgggaatgagatgaagaatatacatgatccaaagt	360
Db	3953	ttttaaattagttgtagcaatgcacatttgggaatgagatgaagaatatacatgatccaaagt	4012
Qy	361	tactcttcacactcaacttcaacatattgttgcacactgtctcccttcaaaagtctgact	420
Db	4013	ttactcttcacactcaacttcaacatattgttgcacactgtctcccttcaaaagtctgact	4072
Qy	421	taacctccaagggagagtaatgcttcggagacctcaagtgaataatccaagaagtggct	480
Db	4073	taacctccaagggagagtaatgcttcggagacctcaagtgaataatccaagaagtggct	4132
Qy	481	gcaagtgtgacttcagaagaacatgaaagtcacagagatgaatactcagggagagtaaatc	540
Db	4133	gcaagtgagacttcagaagaacatgaaagtcacagagatgaatactcagggagagtaaatc	4192
Qy	541	tctgcttaccagcatgtatgtgaaggagttcctcatctccagcagatcaagaibggcatca	600
Db	4193	ttctgcttaccagcatgtatgtgaaggagttcctcatctccagcagatcaagaibggcatca	4252
Qy	601	gtggaactccttttttgaagaatggcaaaagtgttttcaggggaaatcaagactcctt	660
Db	4253	gttggaactccttttttgaagaatggcaaaagtgttttcaggggaaatcaagactcctt	4312
Qy	661	caaacctgtgtgaaactctcagaaccaacogtctactgaactcgtactcctgaaatccac	720
Db	4313	caaacctgtgtgaaactctcagaaccaacogtctactgaactcgtactcctgaaatccac	4372
Qy	721	ccagagttgggtgtcacacagatgtccctcgagagatgagaggttcttggagttggaagcaaga	780
Db	4373	ccagagttgggtgtcacacagatgtccctcgagagatgagaggttcttggagttggaagcaaga	4432
Qy	781	cccttactgaactgagcc 798	
Db	4433	cccttactgaactgagcc 4450	

RESULT 7  
US-08-882-083-1  
; Sequence 1, Application US/08882083  
; Patent No. 5869292  
; GENERAL INFORMATION:  
; APPLICANT: VOORBERG, Johannes J

```

      TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY
      NUMBER OF SEQUENCES: 17
      CORRESPONDENCE ADDRESS:
      ADDRESSEE: Foley & Lardner
      STREET: 3000 K Street, N.W., Suite 500
      CITY: Washington
      STATE: D.C.
      COUNTRY: USA
      ZIP: 20007-5109
      COMPUTER READABLE FORM:
      MEDIUM TYPE: Floppy disk
      COMPUTER: IBM PC compatible
      OPERATING SYSTEM: PC-DOS/MS-DOS
      SOFTWARE: Patentin Release #1.0, Version #1.30
      CURRENT APPLICATION DATA:
      APPLICATION NUMBER: US/08/882,083
      FILING DATE:
      CLASSIFICATION: 514
      PRIOR APPLICATION DATA:
      APPLICATION NUMBER: US 08/558,107
      FILING DATE: 13-NOV-1995
      ATTORNEY/AGENT INFORMATION:
      NAME: ISACSON, John P.
      REGISTRATION NUMBER: 33,715
      REFERENCE/DOCKET NUMBER: 30472/212
      TELECOMMUNICATION INFORMATION:
      TELEPHONE: (202)672-5300
      TELEFAX: (202)672-5399
      TELEX: 904136
      INFORMATION FOR SEQ ID NO: 1:
      SEQUENCE CHARACTERISTICS:
      LENGTH: 5035 base pairs
      TYPE: nucleic acid
      STRANDEDNESS: single
      TOPOLOGY: linear
      FEATURE:
      NAME/KEY: CDS
      LOCATION: 35..5017
      US-08-882-083-1

Query Match      99.0%; Score 790; DB 2; Length 5035;
Best Local Similarity 99.4%; Pred. No. 4.6e-255;
Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 gatcaagtgagatcgtgtggcacaatattattacagcgcataagaccagggtgcgcg 60
DB 4231 GATCAAGTGGATCTGTTGGCACCAATGATTTCACGCGATCAAGACCCAGGTGCCCG 4290
QY 61 tcaagaagttctccagcctctacatctcccaagttatcatcatatgtatagtttgagaa 120
DB 4291 TCAGAAGTTCTCCAGCCTCTACATCTCCAGTTATCATGTAATAGTCTTGATGGGAA 4350
QY 121 gaaagtgagagacttatagaagagatctcaactggaaccttaagtctcttggcaagt 180
DB 4351 GAAGTGGAGAGACTTATCGAGAAATTCACATGGAACCTTAATGCTCTTGGCAATGT 4410
QY 181 ggaattcatctgggataaacaacaatatattttaaaccctcaatatttgctgataacacg 240
DB 4411 GGATTCATCTGGGATAAACAACAATATTTTAAACCTCTCAATTAATTTCTCGAATACATCCG 4470
QY 241 ttgcacccaactcaatataagacattcgagacactctcgatggagttgtagtggcgtga 300
DB 4471 TTTGCACCCAACTCATTTATGACATTCGAGACACTCTTCGATGAGTTGATGGGCTGTGA 4530
QY 301 tttaaatgttcgagcagatgcattgggaaaggagaaagacaatatcagatgcacagat 360
DB 4531 TTTAAATGTTGCAAGCAGCATTGGGAATGAGAGTAAGCAATAATACATGACACAGAT 4590
QY 361 tactgctcaccacttaccataatgatttgccaactggtctccttcaaaagctgact 420
DB 4591 TACTGCTTACCTTACTTATTCACATATGTTGGCACCTGCTGCTTCAAAAGCTGACT 4650

```

```

QY 421 tcactccaagggaggaagtaatctcctggagacactcaagtgaaatccaaagatgct 480
DB 4651 TCACCTCCAAGGAGGAGTAATGCTTGAGACCTCAGGTGAATAATCCAAAGAGTGGCT 4710
QY 481 gcaagtgacttccagaagaacaatgaagaagtcacagaagtaactactcagaagtgaaac 540
DB 4711 GCAAGTGGACTTCCAGAAGCAATGAAGTCAAGAGTAATCTACTCAGAGAGTAAATC 4770
QY 541 tctgcttaacgaatgatagtgaaggagttcctcatctccagcaagtcagaatgaccatca 600
DB 4771 TCTGCTTAACGAGCATGTATGTGAAGAGTCTTCATCTCCACAGATCAAGATGGCATCA 4830
QY 601 gtggactctcttttcaagaatgcaaaagtaagtttctcagggaatcaagactcct 660
DB 4831 GTGACTCTCTTTTTCAGAAATGCAAAAGTAAAGTTTTCAGGGAATAATCAAGACTCCTT 4890
QY 661 cacacctgtgtgaactctctagaaccacggttaactgaactcactcactgaattcacc 720
DB 4891 CACACCTGTGTGACTCTCTCTAGACCCACGTTACTGACTCGCTTCAATTCACACC 4950
QY 721 ccagagttgggtgcacccaagttgcctgaaggatggagttctggcctgcagagcacaga 780
DB 4951 CCAAGTTGGGTGCACCAATGCTCCCTGAGATGGAGTTCTGGGCTCGAGGCACAGCA 5010
QY 781 cctcactgactcgaagc 798
DB 5011 CCTCTACTGAGGGTGGCC 5028

```

```

RESULT 8
US-08-558-107-1
; Sequence 1, Application US/08558107
; Patent No. 5910481
; GENERAL INFORMATION:
; APPLICANT: VOORBERG, Johannes J.
; TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/558,107
; FILING DATE: 13-NOV-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: ISACSON, John P.
; REGISTRATION NUMBER: 33,715
; REFERENCE/DOCKET NUMBER: 30472/212
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5035 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 35..5017
; US-08-558-107-1

```

Query Match 79.0%; Score 790; DB 2; Length 5035;  
 Best Local Similarity 99.4%; Pred. No. 4.6e-255;  
 Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

1 gatcaagtgatgctgttgacccaatgatattcaagcatcaagaccaggtgccc 60  
 4231 GATCAAGTGTGATCTGTGGACCAATGATATTCACGGCATCAAGACCCAGGTCGG 4290  
 Db tcagaagttctcaagctctacatctcagttatcatcatgatagtctgagga 120  
 4291 TCAGAAGTTCTCAGGCTCTACATCTCTAGTTATCATCATGATATAGCTTGAGGGA 4350  
 Db gaagtgacagacttaccaggaattccacctggaaccttaagttctcttggaatg 180  
 4351 GAAGTGCACACTTATCGAGGAATTCACCTGACCTTAATGCTCTCTTGCGCATGT 4410  
 Db ggaattcatctggataaagacaaatatttaacccctcaaatattgctccgataccg 240  
 4411 GGAATTCATCTGGGATTAAMACAAATATTTTAACTCCATTAATGCTCTGATACATCCG 4470  
 Db ttggaccacacatcatatagcatctcagacactcttcgcatgagttgatggctgta 300  
 4471 TTGGACCCACATCATATATAGCATCTGCACACTCTTCGATGAGTGTATGGCTGTGA 4530  
 Db tttaaatagttgacagcatgcatctgggaatgagagtaagacataatcagatgcaagt 360  
 4531 TTTAAATAGTTGACAGCATGCTGCAATGGGAATGAAGATAATACATGATCACAGAT 4590  
 Db taagcttcatctcactcttaccatataatgcttgcacactgctcctcctcaaaactcga 420  
 4591 TACTGCTTCACTACTCTTACCAATATGTTTGCACCTGCTGCTCTCAAACTCGACT 4650  
 Db tcaactcagaaggaggaatgtaagcctcaggaactcaggtgataataatccaaagagtgct 480  
 4651 TCACCTCCAAAGGAGGAGTAATGCTGAGACCTCAGGTGAATTAATCCAAAGAGTGCT 4710  
 Db gcaagtgacttccagaagcaatgaaggtacagagagtaactcactcaggagtaaatc 540  
 4711 GCAAGTGCACACTTCCAGAGCAATGAAAGTACAGAGTAATCTCTCAGGAGTAATATC 4770  
 Db tctgtaccagcatgtaagaggttctcactcactcagcagcatcaadagtgcatca 600  
 4771 TCTGTTACCAAGCATGATGTGAAGAGATTCCTCATCTCCAGCATAGATGCGCATCA 4830  
 Db gtgagctctcttttccagaatgcaagtaaggttttcaagggaatcaagactcctt 660  
 4831 GTGAGCTCTCTTTTTCAGAAATGCAAGTAAAGGTTTTCAGGGAATTCAGACTCCTT 4890  
 Db caccactgtgtaactctcagaccacgcttactgactcgtactcgaattcaacc 720  
 4891 CACACCTGTGTGAGACTCTCTAGACCCGCTTACTGACTGCTGCTTCAATTCACCC 4950  
 Db cagaagttggtgacagagattgcccgtgaagatgaggttcttggtcttgagagagga 780  
 4951 CCAGAGTTGGTGACACCAAGATGCCCCGTGAGATGAGAGGTTCTGGGCTCGAGACAGGA 5010  
 Db ccttactgactcagggc 798  
 5011 CCTTACTGAGGCTGACC 5028

# RESULT 9

US-09-243-539-1

Sequence 1, Application US/09243539

Patent No. 6130203

GENERAL INFORMATION:

APPLICANT: VOORBERG, Johannes J.

TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY

NUMBER OF SEQUENCES: 17

CORRESPONDENCE ADDRESS:

ADDRESSEE: Foley & Lardner

STREET: 3000 K Street, N.W., Suite 500

CITY: Washington

STATE: D.C.  
 COUNTRY: USA  
 ZIP: 20007-5109  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/243,539  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/558,107  
 FILING DATE: 13-NOV-1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: ISACSON, John P.  
 REGISTRATION NUMBER: 33,715  
 REFERENCE/DOCKET NUMBER: 30472/212  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (202)672-5300  
 TELEFAX: (202)672-5399  
 TELEX: 904136  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 5035 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 35..5017  
 US-09-243-539-1

Query Match 79.0%; Score 790; DB 3; Length 5035;  
 Best Local Similarity 99.4%; Pred. No. 4.6e-255;  
 Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

1 gatcaagtgatgctgttgacccaatgatattcaagcatcaagaccaggtgccc 60  
 4231 GATCAAGTGTGATCTGTGGACCAATGATATTCACGGCATCAAGACCCAGGTCGG 4290  
 Db tcagaagttctcaagctctacatctcagttatcatcatgatagtctgagga 120  
 4291 TCAGAAGTTCTCAGGCTCTACATCTCTAGTTATCATCATGATATAGCTTGAGGGA 4350  
 Db gaagtgacagacttaccaggaattccacctggaaccttaagttctcttggaatg 180  
 4351 GAAGTGCACACTTATCGAGGAATTCACCTGACCTTAATGCTCTCTTGCGCATGT 4410  
 Db ggaattcatctggataaagacaaatatttaacccctcaaatattgctccgataccg 240  
 4411 GGAATTCATCTGGGATTAAMACAAATATTTTAACTCCATTAATGCTCTGATACATCCG 4470  
 Db ttggaccacacatcatatagcatctcagacactcttcgcatgagttgatggctgta 300  
 4471 TTGACCCACATCATATATAGCATCTGCACACTCTTCGATGAGTGTATGGCTGTGA 4530  
 Db tttaaatagttgacagcatgcatctgggaatgagagtaagacataatcagatgcaagt 360  
 4531 TTTAAATAGTTGACAGCATGCTGCAATGGGAATGAAGATAATACATGATCACAGAT 4590  
 Db taagcttcatctcactcttaccatataatgcttgcacactgctcctcctcaaaactcga 420  
 4591 TACTGCTTCACTACTCTTACCAATATGTTTGCACCTGCTGCTCTCAAAAGCTCGACT 4650  
 Db tcaactcagaaggaggaatgtaagcctcaggaactcaggtgataataatccaaagagtgct 480  
 4651 TCACCTCCAAAGGAGGAGTAATGCTGAGACCTCAGGTGAATTAATCCAAAGAGTGCT 4710  
 Db gcaagtgacttccagaagcaatgaaggtacagagagtaactcactcaggagtaaatc 540

```

Db 4711 GCAAGTGAAGTCCAGAGACATGAAAGTCCAGAGTACTGAGGAGTAAATC 4770
Qy 541 tctgttaccagatgtagtgaagaggttcctcatctcagcagatcaagaatgcatca 600
Db 4771 TCTGCTTACGACATGATGTGAAGAGTTCCTCATCTCCAGACATCAAGATGGCATCA 4830
Qy 601 gtgagctctcttttcaagaatgcaagaatggttttcaaggaaatcaagactcct 660
Db 4831 GTGAGCTCTCTTTTTCAGATGGCAAGTAAAGTCTTTTCAGGAAATCAACATCTCTT 4890
Qy 661 cacacctgtgtgaactctcagaacccacgttactgactcgcttactcgaatcc 720
Db 4891 CACACCTGTGTGAACCTCTTACACCCAGCTTACTGACTCGCTTCTGAAATTCACCC 4950
Qy 721 ccagaagttgggtgaccagaattgcctgaagatggaagttctgtgctgcagacaga 780
Db 4951 CCAGAGTTGGGTGCACAGATTGCCCTGAGAGATGAGAGTTCTGTGGCTGCAGAGCACAGCA 5010
Qy 781 ccttactgactgagc 798
Db 5011 CCTTACTGAGGTTGCC 5028

```

## RESULT 10

US-08-276-594A-1

```

; Sequence 1, Application US/08276594A
; Patent No. 5693499

```

## GENERAL INFORMATION:

```

APPLICANT: YONEMURA, Hiroshi
APPLICANT: TAJIMA, Yoshitaka
APPLICANT: SUGAWARA, Keishin
APPLICANT: MASUDA, Kenichi
TITLE OF INVENTION: PROCESS FOR PREPARING HUMAN COAGULATION
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/276,594A
FILING DATE: 18-JUL-1994
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JUS 07/950,191
FILING DATE: 24-SEP-1992
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 243262/1991
FILING DATE: 24-SEP-1991
ATTORNEY/AGENT INFORMATION:
NAME: WEGNER, Harold C
REGISTRATION NUMBER: 25,258
REFERENCE/DOCKET NUMBER: 74129/195/AOPA
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136

```

## INFORMATION FOR SEQ ID NO: 1:

```

SEQUENCE CHARACTERISTICS:
LENGTH: 6999 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS

```

```

; LOCATION: 1..6996
US-08-276-594A-1

```

```

Query Match 79.0%; Score 790; DB 1; Length 6999;
Best Local Similarity 100.0%; Pred. No. 5.8e-255;
Matches 790; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 gataaagttgagatctgttggcaccnaatgatatcaccgcatcaagaaccaggtgccg 60
Db 6210 GATCAAGTGGATCTGTGGCACCAATGATTTACGCGCATCAAGACCCAGGTCGCCG 6269
Qy 61 tcaagaattccagcctcacaatcctcaagttatcacaatgatatgcttgatggaa 120
Db 6270 TCAGAAATTTCCAGCCTTACATCTCTCAATGATTAATCAATCAATGATGATGGA 6329
Qy 121 gaagttgcaactatcaggaatccacatgcaactgaactgaactcttcttgcaagtt 180
Db 6330 GAAGTGGCAGACTTATTCAGAGAAATTCACATGAACTTAATGCTTCTTTGCAATCT 6389
Qy 181 gattcactgagataaacaataatlttaacctcaattatgctgatatatccg 240
Db 6390 GATTCATCTGGGATTAACACAAATTTTAAACCTCAATTAATGCTCATATCATCCG 6449
Qy 241 ttgcaaccaactcatatagacatctgcagacactctgcagatgaggttgatggctgga 300
Db 6450 TTGGCAOCCACACTTATATAGCATTTGCGACACTCTTGCATGGAGTGTGAGCTGTGA 6509
Qy 301 tttaaatagttgcaagatgccattgggaatggagaatgaagaatataatgacagat 360
Db 6510 TTTAAATAGTTGAGCATGCCATTTGGGAATGGAGATGAAGAAATATGATGACAGAT 6569
Qy 361 tactgttcatcttacttaccatattgttgcacactgctcctccttaaaagctgact 420
Db 6570 TACTGCTTACTCTTACTTACCAATATGTTGCCACCTGGCTCTCTTCAAAAGCTGCAC 6629
Qy 421 tcaactccaaggagagatgaatgcttggagacatcagtgatataatccaaagatggct 480
Db 6630 TCACCTCCAAAGGAGAGATATGCTGGAGACTCAGGTGATATATCCAAAGAGTGGCT 6689
Qy 481 gaaagtggactcccaagaagacaaatgaagtcaagagtaactcaaggagtaaatc 540
Db 6690 GCAAGTGCATTCGAAAGACAAATGAAGTCAAGAGTAACTCAAGGAGTAATATC 6749
Qy 541 tctgttaccagatgtagtgaagaggttcctcatctcagcagatcaagaatgcatca 600
Db 6750 TCTGCTTACGACATGATGTGAAGAGTTCCTCATCTCCAGACATCAAGATGGCATCA 6809
Qy 601 gtgagctctcttttcaagaatgcaagaatggttttcaaggaaatcaagactcct 660
Db 6810 GTGAGCTCTCTTTTTCAGATGGCAAGTAAAGTCTTTTCAGGAAATCAACATCTCT 6869
Qy 661 cacacctgtgtgaactctcagaacccacgttactgactcgcttactcgaatcc 720
Db 6870 CACACCTGTGTGAACCTCTTACACCCAGCTTACTGACTCGCTTCTGAAATTCACCC 6929
Qy 721 ccagaagttgggtgaccagaattgcctgaagatggaagttctgtgctgcagacaga 780
Db 6930 CCAGAGTTGGGTGCACAGATTGCCCTGAGAGATGAGAGTTCTGGGCTGCAGGACACAGCA 6989
Qy 781 ccttactgag 790
Db 6990 CCTTACTGA 6999

```

## RESULT 11

US-08-121-202-1

```

; Sequence 1, Application US/08121202
; Patent No. 5563045

```

## GENERAL INFORMATION:

```

APPLICANT: Pitman, Debra
APPLICANT: Rehmetulla, Alnawaz
APPLICANT: Wozney, John M.

```

APPLICANT: Kaufman, Randal J.  
TITLE OF INVENTION: CHIMERIC PROCOAGULANT PROTEINS  
NUMBER OF SEQUENCES: 27  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Legal Affairs, Genetics Institute, Inc.  
STREET: 87 Cambridgepark Drive  
CITY: Cambridge  
STATE: MA  
COUNTRY: USA  
ZIP: 02140  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/121,202  
FILING DATE: 14-SEP-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Melner, M. C.  
REGISTRATION NUMBER: 31,544  
REFERENCE/DOCKET NUMBER: GI 5195A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 876-1210 X8574  
TELEFAX: (617) 876-5851  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7056 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..7053  
US-08-121-202-1

Query Match 79.08; Score 790; DB 1; Length 7056;  
Best Local Similarity 100.0%; Pred. No. 5.8e-255;  
Matches 790; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 gatcaaggtagatcgttggcaccatattatcagcagcaatcgaagccagggatgccg 60  
|||  
Db 6267 GATCAAGGTGATCTGTTGGACCAATGATTATTCAGCGATCAAGACCAGGGTGCCTG 6326  
Oy 61 tcagaagttccagcctctcatctcagttatcatcatatgtagtcttgaggaa 120  
|||  
Db 6327 TCAGAAGTTCCAGCCCTCTACATCTCATCTTATCATCATGATGATCTGTGATGGAA 6386  
Oy 121 gaagtggcagactatcgaggaatctccactggaacttaagtctcttggcaatgt 180  
|||  
Db 6387 GAAGTGGCAGACTATTCAGGAAATTCACATGAACTTAATGCTCTTGGCAATGT 6446  
Oy 181 ggattcacttggaataaacaacataattttaaccctcccaatlaibctcgatacatcg 240  
|||  
Db 6447 GGATTCACTTGGAATAAACAATATTTTAACCTCCCAATTTATGCTGCATACATCCG 6506  
Oy 241 ttggacccaactcaatataagcattcgacactcttcgacatgagtgatggctgtga 300  
|||  
Db 6507 TTTGACCCCACTCATTTATGACATTCGACGACCTTTCGATGAGTGTGATGGCTGTGA 6566  
Oy 301 ttttaagtgtgagcagatgcatatgggahbgagaaagcaaatatcatagatgcaagat 360  
|||  
Db 6567 TTTAATAGTTGACGACATGCCATTGGGAATGAGATAAAGCAATATCAATGACACAGAT 6626  
Oy 361 tactgtctcatccacttaccatataatgtttgacactgtgtcctccaaagtgcact 420  
|||  
Db 6627 TACGTCTTCATCTTACTTACCATATGTTTGGCACTGCTGCTTCAAAAGCTGCACT 6686

Oy 421 tcacctcagaaggaggaatgctctgagacctcaagtgaataatccaaagtgcact 480  
|||  
Db 6687 TCACCTCAGAGGAGGAGTATGCTTGAGACCTCAGCTCAATATATCCAAAGATGCT 6746  
Oy 481 gcaagtgaacttccagaagaacatgaaatgacagaggaactaactcgaaggagtaaatc 540  
|||  
Db 6747 GCAAGTGAAGTCCAGAAACAAATGAAGTACAGAGGAGTAACTACAGGAGTAAATC 6806  
Oy 541 tctgtaccagatgatatgaaaggagttccctcatctccagcaagtcagatgcatca 600  
|||  
Db 6807 TCTGTTACAGCATGTATGTGAAGGAGTTCCTCATCTCCAGAGTCAAGATGGCCATCA 6866  
Oy 601 gtgacactcttcttcaagaatgcaagaaggttttcaaggaaatcaaacctc 660  
|||  
Db 6867 GTGACCTCTTTTTCAGATGCGAAAGTAAAGTTTTCAGGAAATCAACACTCCTT 6926  
Oy 661 caacactgtgtgaactctctagaccaccgttactgactgctactccttgaatcaacc 720  
|||  
Db 6927 CACACGTGTGTAACCTCTAGACCCACCGCTTACTGACTGCTACTCTTCGATTCACCC 6986  
Oy 721 ccagagttggtgacacagatgcccctgagagatgaggttctggtctgagagcaaga 780  
|||  
Db 6987 CCAGAGTTGGTGCACCAAGTGCCTGAGAGATGAGGTTTGGCTGCGAGGACACAGA 7046  
Oy 781 cctctactga 790  
|||  
Db 7047 CCTCTACTGA 7056

RESULT 12  
US-08-366-851A-1  
Sequence 1, Application US/08366851A  
Patent No. 5681746

## GENERAL INFORMATION:

APPLICANT: Bodner, Mordechai  
APPLICANT: De Polo, Nicolas J.  
APPLICANT: Hsu, David Chi-Tang  
APPLICANT: Chang, Steven  
TITLE OF INVENTION: Retroviral Delivery of Full Length Factor VIII  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Viagene, Inc.  
STREET: 11055 Roselle Street  
CITY: San Diego  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 92121

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/366,851A  
FILING DATE:

## CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:  
NAME: Chambers, Daniel M.  
REGISTRATION NUMBER: 34,561  
REFERENCE/DOCKET NUMBER: 930049.438  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 452-1288  
TELEFAX: (619) 452-2616  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:

LENGTH: 8967 base pairs

TYPE: nucleic acid

STRANDEDNESS: both

TOPOLOGY: unknown

MOLECULE TYPE: CDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 110..7165

US-08-366-851A-1

Query Match 79.0%; Score 790; DB 1; Length 8967;  
 Best Local Similarity 99.4%; Pred. No. 6.8e-255;  
 Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 gatcaagagtgatctgttgcccaatgatattatccagcatcaagaaccaggtcccg 60  
 DB 6376 GATCAAGAGTGTGATCTGTGGCCATGATGATTTACAGGCATCAACAGGATCCCG 6435  
 QY 61 tcagaagttccagcctctacatctccagttatcatcatatgtagtcttgatggaa 120  
 DB 6436 TCAGAAGTTCAGGCTCAGATCTCAGTTATCATCATCATATAGTCTTATGGAA 6495  
 QY 121 gaagtgagacttalcagagaatctccatctgaaaccttaatgcttcttggcaatgt 180  
 DB 6496 GAAGTGCAGACTTATTCAGAGAAATTCACCTGAACCTTAATGCTCTTGGCAATGT 6555  
 QY 181 ggaattcatctggagataaacaatatttttaaccctccaatattgtctcgatataccg 240  
 DB 6556 GGATTCATCTGGGATTAACACATATTTTAACTCCATATTATGCTCGATACATCCG 6615  
 QY 241 ttggacccaactcatatagcatctcgacgactcttcgcatgagttgagtggtctgta 300  
 DB 6616 TTGGACCCCAACTCATATATAGCATCTTCGACACTCTTCGATGAGTGTGAGGCTGTGA 6675  
 QY 301 tttaaatagtctgacgatccatctgggaatggagagtaagaacatctcgatgacagat 360  
 DB 6676 TTTAAATAGTTGGAGCATGTCGTAATGGGAATGAGAGTAAGCAATATCATGATCACAGAT 6735  
 QY 361 tactgcttcatcttacttaccatatagtttgacacctgtctcccttcaaaaagctgact 420  
 DB 6736 TACTGCTTACTCTACTTTACCAATATGTTTGGCACCTGTCTCTTCAAAAAGCTGACT 6795  
 QY 421 tcaactcaaggagaggaatgtaagtcctgaggaactcaggtgaataatccaaaagtgtgct 480  
 DB 6796 TCACCTCCAGGAGGAGGATGATGCTGAGAGCCTCAGTGAATTAATCCAAAAGCTGCT 6855  
 QY 481 gcaagtgacttccagaagaacatgaaatcacagagagtaacacccaagaggaataatc 540  
 DB 6856 GCAAGTGGACTTCCAGAGACATGAAGTCAACAGAGTACTACTCAGGAGGATTAATC 6915  
 QY 541 tctgcttaccagatgtagtgaagagttcctcatctccagcagtaagaatggcatca 600  
 DB 6916 TCTGCTTACAGCATGATGATGAGAGTCTCATCTCCAGCAGTCAAGATGGCATCA 6975  
 QY 601 gtggactctcttttttcaagaatggcaagtaaggttttccagggaaatcaaacctctt 660  
 DB 6976 GTGGACTCTCTTTTTCAGATGAGCAAGTAAGGTTTTCAGGGAATCAAGACTCTCTT 7035  
 QY 661 caacccgtgtgaactctctagaccacacgttactgactcgttacctcgaattcaacc 720  
 DB 7036 CACACCTGTGTGAATCTCTAGACCCACCTGACTACTGCTACTTGAATTCACCC 7095  
 QY 721 ccagaagttgggtgcacacagatctccctgaagatgaggtctctggtctgcaagacagaa 780  
 DB 7096 CCAGAAGTGGGTGCACACAGATTTGGCCCTGAGAGTGAAGGTTCTGGGCTGCAGGACAGGA 7155  
 QY 781 cctctactgactcgagcc 798  
 DB 7156 CCTCTACTGAGGGGTGGCC 7173

RESULT 13

US-07-864-004B-3  
 Sequence 3, Application US/07864004B  
 Patent No. 5364771

GENERAL INFORMATION:

APPLICANT: Iollar, John S.  
 APPLICANT: Runge, Marschall S.

TITLE OF INVENTION: Hybrid Human/Porcine Factor VIII  
 NUMBER OF SEQUENCES: 6

## CORRESPONDENCE ADDRESS:

ADDRESSEE: Kilpatrick & Cody  
 STREET: 1100 Peachtree Street  
 CITY: Atlanta  
 STATE: Georgia  
 COUNTRY: US

ZIP: 30309

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/864,004B

FILING DATE: 07 APRIL 1992

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Padst, Patricia L.

REGISTRATION NUMBER: 31,284

REFERENCE/DOCKET NUMBER: EMU106

TELECOMMUNICATION INFORMATION:

TELEPHONE: 404-815-6508

TELEFAX: 404-815-6555

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 9009 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: N-terminal

ORIGINAL SOURCE:

ORGANISM: Homo sapien

TISSUE TYPE: Liver

FEATURE:

NAME/KEY: misc-feature (Domain Structure)

LOCATION: 1...2277

OTHER INFORMATION: /note= "Equivalent to the A3-C1-C2"

OTHER INFORMATION: domain"

US-07-864-004B-3

Query Match 79.0%; Score 790; DB 1; Length 9009;  
 Best Local Similarity 99.4%; Pred. No. 6.8e-255;  
 Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 gatcaagagtgatctgttgcccaatgatattatccagcatcaagaaccaggtcccg 60  
 DB 6417 GATCAAGAGTGTGATCTGTGGCCATGATGATTTACAGGCATCAACAGGATCCCG 6476  
 QY 61 tcagaagttccagcctctacatctccagttatcatcatatgtagtcttgatggaa 120  
 DB 6477 TCAGAAGTTCAGGCTCAGATCTCAGTTATCATCATCATATAGTCTTATGGAA 6536  
 QY 121 gaagtgagacttalcagagaatctccatctgaaaccttaatgcttcttggcaatgt 180  
 DB 6537 GAAGTGCAGACTTATTCAGAGAAATTCACCTGAACCTTAATGCTCTTGGCAATGT 6596  
 QY 181 ggaattcatctggagataaacaatatttttaaccctccaatattgtctcgatataccg 240  
 DB 6597 GGATTCATCTGGGATTAACACATATTTTAACTCCATATTATGCTCGATACATCCG 6656  
 QY 241 ttggacccaactcatatagcatctcgacgactcttcgcatgagttgagtggtctgta 300  
 DB 6657 TTGGACCCCAACTCATATATAGCATCTTCGACACTCTTCGATGAGTGTGAGGCTGTGA 6716



Query Match	Best Local Similarity	Matches	79.0%	Score 790;	DB 1:	Length 9009;
			99.4%;	Pred. No. 6.8e-255;		
			Conservative 0;	Mismatches 5;	Indels 0;	Gaps 0;
US-08-251-937A-3						
SEQUENCE CHARACTERISTICS:						
LENGTH: 9009 base pairs						
TYPE: nucleic acid						
STRANDEDNESS: single						
TOPOLOGY: linear						
MOLECULE TYPE: cDNA						
HYPOTHETICAL: NO						
ANTI-SENSE: NO						
FRAGMENT TYPE: N-terminal						
ORIGINAL SOURCE:						
ORGANISM: Homo sapien						
TISSUE TYPE: Liver						
FEATURE:						
NAME/KEY: misc_feature (Domain Structure)						
LOCATION: 5001..7053						
OTHER INFORMATION: /note="Equivalent to the A3-C1-C2						
OTHER INFORMATION: /note="Equivalent to the A1-A2						
OTHER INFORMATION: domain"						
NAME/KEY: misc_feature (Domain Structure)						
LOCATION: 1...2277						
OTHER INFORMATION: /note="Equivalent to the A1-A2						
OTHER INFORMATION: domain"						
US-08-251-937A-3						
Query Match						
Best Local Similarity						
Matches 793; Conservative 0;						
US-08-251-937A-3						
1 gatcaaggtgagctcgtttggcacaatgatattcaagcgcatccaagcaggtgccg 60						
6417 GATCAAGGTGAGCTCGTTGGCACCAATGATTTTCACGGCATCAAGCCAGGAGTCCCG 6476						
61 tcagaagttcttcacagcctctacatctctcagittatcatcalgtalagtctgatggaa 120						
6477 TCAGAAGTTCTTCACAGCCTCTACATCTCTCAGTTTATCATCATGTAATGCTGATGGAA 6536						
121 gaagtggcagactatcgaagaaatccacatcgtgaaccttaatggctctttggcaagt 180						
6537 GAAGTGGCAGACTTATCGAGAAATTCACCTGGAACCTTAATGGCTCTTTGGCAATGT 6596						
181 gattcatcttggtgataaacaacataattttaacctccaattatgtgcgatataccg 240						
6597 GATTTCATCTGGGATTAACACATAATTTTAAACCTCCAAATATTTGCTCGATACATCCG 6656						
241 ttgcaccacaactcatatagcatctgcagacactcttcgacatgaggttgaatggctgtga 300						
6657 TTTCGACCCAACTCATTAATAGCATTCGACGACACTCTTCGATGGAGTTGATGGGCTGTGA 6716						
301 tttaaatagttgcagcatgcatctgggaatggagagtaagcaataatcagatgcacagat 360						
6717 TTTAATAGTTGTCACACATGCCCCATTTGGGAATGAGATTAAGCAATATCAGATGCACAGAT 6776						
361 tactcttcacatctacttaccaatatggtttgcacctgggtcttcctttaaagatcagat 420						
6777 TACTCTTCATCTTACTTTACCAATATGTTTCCACCTGGCTCTCTTCAAAAGCTTGACT 6836						
421 tcaactccaagagagagtaatgaccttgcagacactcagtgtaaataatccaaaagatggct 480						
6837 TCACCTCCCAAGGAGAGATGATGCCGTGGAGACTCTAAGTGTAATCCAAAAGATGGCT 6896						
481 gcaagtggacttccagaagaatgaagatccaagagtaactactcaggagtaaatc 540						
6897 GCAAGTGGACTTCCAGAGACAAATGAATGCAAGAGTAATACTACAGGAGTAATAATC 6956						
541 tctgcttaccagcatgatagtgaaggagttctcaatctccaagatccaagatggcaca 600						
6957 TCTGCTTACACACATGATGTAAGAGTCTCTCATCTCCACACAGTCAAGATGGCCATCA 7016						
601 gtggactcctcttttccaagaatggcaaatgaagtttttcaggagaaatccaagactcctt 660						
7017 GTGGACTCTCTTTTTCAGAAATGGCAAAATGAAGTTTTCAGGGGAATCAAGACTCTTT 7076						

QY 661 cacacctgtggtgacctctagaccaccgcttactgacctgacctgcaattccccc 720  
|||||  
Db 7077 CACACCTGTGGTGAACCTCTAGACCCACCTACTGACTGCTACCTGCAATTCACACC 7136  
QY 721 ccagagttgggtgacccagatgcccctgagagatgaggttctggcgcgagagacagga 780  
|||||  
Db 7137 CCAGAGTTGGGTGACACCGATGATGCCCTGAGAGTGTGCGCTGCGAGCACAGGA 7196  
QY 781 cctctactgactcgagcc 798  
|||||  
Db 7197 CCTCTACTGAGGTGGCC 7214  
RESULT 15  
US-08-212-133A-1  
Sequence 1, Application US/08212133A  
Patent No. 5663060  
GENERAL INFORMATION:  
APPLICANT: Lollar, John S.  
APPLICANT: Runge, Marschal S.  
TITLE OF INVENTION: Hybrid Human/Animal Factor VIII  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Kilpatrick & Cody  
STREET: 100 Peachtree Street  
CITY: Atlanta  
STATE: Georgia  
COUNTRY: US  
ZIP: 30303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/212,133A  
FILING DATE: March 11, 1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/864,004  
FILING DATE: 07-Apr-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Pabst, Patricia L.  
REGISTRATION NUMBER: 31,284  
REFERENCE/DOCKET NUMBER: EMU/76677  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 404-572-6508  
TELEFAX: 404-572-6555  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9009 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Homo sapien  
TISSUE TYPE: Liver  
FEATURE:  
NAME/KEY: misc\_feature (Domain Structure)  
LOCATION: 5125..7053  
OTHER INFORMATION: /note= "Equivalent to the A3-C1-C2  
OTHER INFORMATION: domain"  
FEATURE:  
NAME/KEY: misc\_feature (Domain Structure)  
LOCATION: 1..227  
OTHER INFORMATION: /note= "Equivalent to the A1-A2 domain."  
FEATURE:  
NAME/KEY: Domain  
LOCATION: 1..227  
OTHER INFORMATION: /note= "cDNA encoding human factor

OTHER INFORMATION: VIII.  
US-08-212-133A-1  
Query Match 79.0%; Score 790; DB 1; Length 9009;  
Best Local Similarity 99.4%; Pred. No. 6.8e-255;  
Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 1 gatcaagtgatctgttggcaccatgtatttaccgcatcgaagaccagagtgcccg 60  
Db 6417 GATCAAGTGTGATCTGTGGCACCATGATTTATCCGCGATCAGACCCAGAGTCCCG 6476  
QY 61 tcagaagttctccagcctctacatctctcagttatcaltcatgtaagttcgtatgggaa 120  
Db 6477 TCAGAAGTTCTCCAGCCTCTACATCTCTCAGTTTATCATCATGATGATAGCTTGATGGGAA 6536  
QY 121 gaagtggcagacttaccgaggaattccacttgaaccttaatggtcttcttggcaagt 180  
Db 6537 GAAGTGGCAGACTTATCGAGGAATTCACCTGGAACCTTAATGCTTCTTTGGCAAGT 6596  
QY 181 ggaatcatctggagataaacaacatatctttaaacctccaattatgtctgatatccg 240  
Db 6597 GGATTCATCTGGGATTAACACATATTTTAACTCCATATTATTCCTGATACATCCG 6656  
QY 241 ttgcaaccaactcaltatagatctgcagcactcttcgcatgagttgatlgtgcttga 300  
Db 6657 TTGCAACCAACTCATTTATAGATTTGCGACACTCTTCGATGAGATTGATGGCTGTGA 6716  
QY 301 tttaaatagttgcagcatgccatgtggaaatggaatgaagcaatatcagatgcagat 360  
Db 6717 TTTAAATAGTTGCAGCATGCCATTTGGGAATGAGATAAGCAATATCGATGCACAGAT 6776  
QY 361 taatgttcaatctacttaccatatgtttgcacactgtctcctcaaaagcccgat 420  
Db 6777 TACTGCTTACTCTACTTTTACCAATATGTTTGCACACTGTCTCTCTTCAAAAGCTCGACT 6836  
QY 421 tcacctccaaggaggaagtaatgctcggagacctcagtgtaataatccaaaagtggt 480  
Db 6837 TCACCTCCAAGGAGGAGTAATGCTCGAGACTCGATGATTAATCCAAAGAGTGGCT 6896  
QY 481 gcaagtgagcttcgaagaacatgaagtcagaggaagtaactactcggggtaaatc 540  
Db 6897 GCAAGTGAAGCTTCAGAGAACAATGAAGTCAAGAGTAAGTAAGTCAAGGAGTAAGT 6956  
QY 541 tctgcttaacagatgcatgtaagagagttcctcactcctcaagcagtcgaagtgcacatca 600  
Db 6957 TCTGCTTACAGCATGTATGTGAAGAGTTCCTCATCTCCACAGATCAAGATGGCCATCA 7016  
QY 601 gtgagctctcttttccagaatggaagtaagtttccagggaatcaagactcctt 660  
Db 7017 GTGAGCTCTCTTTTTCAGAAATGGCAAGTAAGTAAAGTTTTCAGGGAAATCAAGACTCCTT 7076  
QY 661 cacacctgtggtgacctctagaccaccgcttactgacctgacctgcaattccccc 720  
Db 7077 CACACCTGTGGTGAACCTCTAGACCCACCTACTGACTGCTACCTGCAATTCACACC 7136  
QY 721 ccagagttgggtgacccagatgcccctgagagatgaggttctggcgcgagagacagga 780  
Db 7137 CCAGAGTTGGGTGACACCGATGATGCCCTGAGAGTGTGCGCTGCGAGCACAGGA 7196  
QY 781 cctctactgactcgagcc 798  
|||||  
Db 7197 CCTCTACTGAGGTGGCC 7214

Search completed: January 8, 2002, 17:25:52  
Job time: 12603 sec